

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw. model

Run on: April 29, 2001, 06:40:29 ; Search time 14.58 Seconds

(without alignments)  
608.042 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678

Sequence: 1 MSPPKYPSSLRTPETLPD.....DRKKLIQEGKLDRTFLSY 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	2 JE0383	NADH dehydrogenase
2	514	75.8	129	2 S28237	NADH dehydrogenase
3	315	46.5	133	2 JC2003	NADH ubiquinone ox
4	149	22.0	179	2 A33164	hypothetical prote
7	88.5	13.1	1078	2 S77162	DNA topoisomerase
7	74	10.9	308	2 G83104	probable ferredoxi
7	74	10.9	451	2 F75177	tryptophan synthas
8	70.5	10.4	788	2 S75209	dark protein - Syn
9	70	10.3	925	2 G64327	H+-transporting AT
10	70	10.3	1025	2 T18376	multidrug resistan
11	69.5	10.3	294	2 S22440	protein kinase (EC
12	69	10.2	298	2 S69523	hypothetical prote
13	69	10.2	1023	2 S12519	glutathionin - fruit
14	69	10.2	1660	2 T18561	vitellinogen vit-6
15	68.5	10.1	294	2 S23095	protein kinase (EC
16	68.5	10.1	294	2 T49271	CELL DIVISION CONT
17	68.5	10.1	574	2 S57072	hypothetical prote
18	68.5	10.1	1330	2 A36373	hypothetical prote
19	68.5	10.1	1333	2 T38401	retrotransposabl
20	68	10.0	185	2 S12205	hypothetical prote
21	68	10.0	716	1 JC5061	macrophage-stimula
22	67.5	10.0	133	1 MNVQBY	genome-linked prot
23	67.5	10.0	289	2 E75391	conserved hypothet
24	67.5	10.0	296	2 C82645	transcription regu
25	67.5	10.0	600	2 C83142	hypothetical prote
26	67	9.9	194	2 A82072	conserved hypothet
27	67	9.9	232	2 B75121	hypothetical prote
28	67	9.9	316	2 T21180	hypothetical prote
29	67	9.9	455	2 S71344	purinergic recepto

30	66.5	9.8	472	2 B37777	methyl viologen-re
31	66.5	9.8	674	2 E83263	hypothetical prote
32	66.5	9.8	634	2 S32230	Ca2+-transporting
33	66.5	9.8	1036	2 A57386	preprotein translo
34	66	9.7	193	2 H81874	probable outer mem
35	66	9.7	271	2 T24965	hypothetical prote
36	66	9.7	342	2 D71324	conserved hypothet
37	66	9.7	353	2 F83570	hypothetical prote
38	66	9.7	413	2 H75357	tRNA (5-methylamin
39	66	9.7	657	1 F0V2EV	major core protein
40	66	9.7	677	2 S54561	RNA14 protein - ye
41	65.5	9.7	354	2 T22108	hypothetical prote
42	65.5	9.7	899	2 B38529	nix protein - Bsc
43	65.5	9.7	959	2 D83339	glycine cleavage s
44	65.5	9.7	2358	2 T39569	probable alpha-glu
45	65.5	9.7	2371	2 T43432	alpha-glucan synth

## ALIGNMENTS

RESULT 1  
JE0383  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NDUFB4 - human  
N:Alternate names: NADH:ubiquinone oxidoreductase subunit NDUFB4  
C:Species: Homo sapiens (man)  
C>Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 21-Jul-2000  
R:Loeffen, J.L.C.M.; Triepels, R.H.; van den Heuvel, L.P.; Schuelke, M.; Buskens, C.A  
Biochem. Biophys. Res. Commun. 253, 415-422, 1998  
A>Title: cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: Hu  
A:Reference number: JE0379; MUID:99097250  
A:Accession: JE0383  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-129 <LOE>  
A:Cross-References: NID:94164445; PIDN:AAD05421.1; PID:94164446  
C:Keywords: NAD; oxidoreductase

Query Match 100.0%; Score 678; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 6; le-65;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MSPPKYPSSLRTPETLPDAEYNISPEPTRAOAERLATRAQKREYLLQYNDPNRGLI 60  
DB 1 MSPPKYPSSLRTPETLPDAEYNISPEPTRAOAERLATRAQKREYLLQYNDPNRGLI 60  
OY 61 ENPALLRMAYARTINYPNFRPTPKNSLMGALCGFGLIFYYIIKTERDRKKLEQEGK 120  
DB 61 ENPALLRMAYARTINYPNFRPTPKNSLMGALCGFGLIFYYIIKTERDRKKLEQEGK 120  
OY 121 LDRTFHLSY 129  
DB 121 LDRTFHLSY 129  
RESULT 2  
S28237  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain B15 - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
R:Walker, J.E.; Arizumi, J.M.; Dupuis, A.; Fearnsley, I.M.; Finel, M.; Medd, S.M.; P  
J. Mol. Biol. 226, 1051-1072, 1992  
A>Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine hear  
A:Reference number: S28237; MUID:92389317  
A:Accession: S28237  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-129 <WAL>  
A:Cross-References: EMBL:X64898; NID:94113; PIDN:CAA46107.1; PID:94114  
C:Keywords: electron transfer; mitochondrion; NAD; oxidoreductase

Query Match 75.8%; Score 514; DB 2; Length 129;  
 Best Local Similarity 73.6%; Pred. No. 1.7e-47;  
 Matches 93; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

OY 1 MSFPPKPSLTLPETLDPAEYNISPTERRAQAERLAIRAOLKREYLLQYNDPNRGLI 60  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 1 MSFPPKPSLTLPETLDPAEYDISSETRKAQAERLAIRSLKREYOLQYDPSRGGVI 60  
 OY 61 ENPALIKRAVARTINVPNFRPTPKNSLMGALCGFLIFYYIKTERDREKELIOEGK 120  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 61 EDPALVARTVARSANIVNFRPTPKNSLMGALFGIGPLVFWYVEKTDREKREKLIQEGK 120  
 OY 121 LDRPFLSY 129  
 ||||| : : : : :  
 Db 121 LDRPFLSY 129

RES 3

JC  
 NADH ubiquinone oxidoreductase B15 chain like protein - chicken  
 N:Alternate names: murine Hox-3.1 homeobox  
 C:Species: Gallus gallus (chicken)  
 C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 07-Feb-1997  
 C:Accession: J02003  
 R:Goldberg, G.S.; Kaczmarczyk, W.  
 Gene 133, 233-235, 1993  
 A:Title: A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1 homeobox  
 A:Reference number: J02003; MUID:94040816  
 A:Accession: J02003  
 A:Molecule type: DNA  
 A:Residues: 1-133 <GOL>  
 C:Genetics:  
 A:Gene: g9hpw  
 A:Introns: 67/3  
 C:Keywords: homeobox; phosphoprotein; sulfoprotein; transmembrane protein  
 F:95-112/Domain: transmembrane #status predicted <TMM>  
 F:30/Binding site: sulfate (Tyr) (covalent) #status predicted  
 F:114/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 46.5%; Score 315; DB 2; Length 133;  
 Best Local Similarity 59.0%; Pred. No. 2.5e-26;  
 Matches 62; Conservative 14; Mismatches 29; Indels 0; Gaps 0;

OY 5 KYKPSLTLPETLDPAEYNISPTERRAQAERLAIRAOLKREYLLQYNDPNRGLIENPA 64  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 12 EYRPNRYVSLPAELDPATYDPLEKRAEAERLAIRARLQYLLQNTPKPPVIEDPA 71  
 OY 65 LLRMVARTINVPNFRPTPKNSLMGALCGFLIFYYIKTER 109  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 72 LLRMVARTINVPNFRPTPKNSLMGALCGFLIFYYIKTER 116

RESULT 4

A33164  
 hypothetical protein walter - chicken  
 N:Alternate names: homeotic protein Hox-3.1 homolog; hypothetical 13.1k protein  
 C:Species: Gallus gallus (chicken)  
 C:Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 21-Jul-2000  
 C:Accession: A33164; J01395  
 R:Goldberg, G.S.  
 submitted to the Protein Sequence Database, March 1991  
 A:Reference number: A33164  
 A:Accession: A33164  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-179 <GOL>  
 A:Experimental source: strain Leghorn  
 R:Goldberg, G.S.; Kaczmarczyk, W.  
 Gene 121, 397-398, 1992  
 A:Title: Sequence of a novel chicken genomic DNA fragment that hybridizes to the murine

A:Reference number: J01395; MUID:93077061  
 A:Accession: J01395  
 A:Molecule type: DNA  
 A:Residues: 1-124 <GOL>  
 A:Cross-references: GB:M84354; NID:g211948; PIDN:AAA70193.1; PID:g211949

Query Match 22.0%; Score 149; DB 2; Length 179;  
 Best Local Similarity 60.0%; Pred. No. 1.6e-08;  
 Matches 30; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

OY 5 KYKPSLTLPETLDPAEYNISPTERRAQAERLAIRAOLKREYLLQYNDP 54  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 26 EYRPNRYVSLPAELDPATYDPLEKRAEAERLAIRARLQYLLQNTLP 75

RESULT 5

S77162  
 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Synechocystis sp. (strain  
 N:Alternate names: DNA gyrase chain B; protein s112005  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S77162  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, Y.;  
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Matsubae, A.; Yamada, M.; Yas  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
 S.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S77162  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1078 <KAN>  
 A:Cross-references: EMBL:D90908; GB:AE001339; NID:91652725; PIDN:BA17720.1; PID:d101  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Gene: gyrB  
 A:Start codon: GTG  
 C:Keywords: isomerase

Query Match 13.1%; Score 88.5; DB 2; Length 1078;  
 Best Local Similarity 28.2%; Pred. No. 0.4;  
 Matches 33; Conservative 13; Mismatches 42; Indels 29; Gaps 4;

OY 22 EYNISPTERRAQAERL-----AIRAOLKREYLLQYNDPNRGLIENPALIKRAVART 73  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 704 EYMAQAKNRTAQAERVRQHFQAPNGLRQOYSENAAVKNW-----NPETLKWRQKT 754  
 OY 74 INVY-PNFRPTPKNSLMGALCGFLIFYYIKTERDREKELIOEGKLDRTFHSY 129  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 755 KEQWTFERREKRELAQT-----YRKTLALQVEIENGVLIDISAYDSY 800

RESULT 6

G83104  
 probable ferredoxin reductase PA4331 [imported] - Pseudomonas aeruginosa (strain PA01  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: G83104  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
 Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: G83104  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-308 <STO>  
 A:Cross-references: GB:AE004849; GB:AE004091; NID:99505050; PIDN:AG07719.1; GSPDB:GN  
 A:Experimental source: strain PA01

C:Genetics:  
A:Gene: PA4331

Query Match 10.9%; Score 74; DB 2; Length 308;  
Best Local Similarity 35.7%; Pred. No. 2.9;  
Matches 20; Conservative 5; Mismatches 23; Indels 8; Gaps 1;

QY 13 TLPLETPAEVNIISPETRAQAERLAIRAOLKREYLLOYNDPNRGLIENPALLRW 68  
DB 54 TLPLEALDP-----VRREAGMRLAACCCRVGLDLVLOPFDEPRDGLPARVYACWH 101

RESULT 7  
F75177

tryptophan synthase, chain beta (trpb-2) PAB1970 - *Pyrococcus abyssi* (strain Orsay)  
C:Species: *Pyrococcus abyssi*  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: F75177

R:Annotation: Genoscope

A:Description: to the EMBL Data Library, July 1999

A:Reference number: A75001

A:Accession: F75177

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 <KAM>

A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457330; PIDN:CAB49501.1; PID:g545801

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1970

C:Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology

F:29-432/Domain: tryptophan synthase beta chain homology <S>

Query Match 10.9%; Score 74; DB 2; Length 451;

Best Local Similarity 32.6%; Pred. No. 4.7;

Matches 29; Conservative 10; Mismatches 26; Indels 24; Gaps 5;

QY 11 LRTLEPTLDPAEVNIISPET-----RRQAERLAIRAOLKREYLLOYNDPNRGLIE 61  
DB 29 LLDLPLEPLDP---LDPEEPEPIDIEKLRIFAEL-VKQELSR--RY-----IE 73  
QY 62 NPALLMAYARTINYPNRPPTPKNSLMG 90  
DB 74 IPEGLRLKLYSKIGRPPLFRATNLEKLLG 102

RF 8

S7 dnak protein - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein sll1932

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S75209

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S75209

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-788 <KAM>

A:Cross-references: EMBL:D90903; GB:AB001339; NID:g1652127; PIDN:BAAL7123.1; PID:g165219

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: dnak

A:Start codon: GTG

C:Superfamily: heat shock protein 70

Query Match 10.4%; Score 70.5; DB 2; Length 788;  
Best Local Similarity 30.5%; Pred. No. 22;  
Matches 25; Conservative 9; Mismatches 33; Indels 15; Gaps 3;

QY 16 ETPLDPAEVNIISPETRAQAERLAIRAOLKREYLLOYNDPNRGLIENPALLRWARTI- 74  
DB 587 EILDSLEKDERLDRQAADLDVLYELNREVLQYDD-KEEGFE-----AIKTFE 638  
QY 75 -----NYPNRPPTPKNSLMG 90  
DB 639 GDFDDDDYNNRRPAPRDYRG 660

RESULT 9  
G64327

H-transferring ATP synthase (EC 3.6.1.34) chain I - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: G64327

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A:Reference number: A64300; MUID:96337999

A:Accession: G64327

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-695 <BUJ>

A:Cross-references: GB:U67478; GB:L77117; NID:g1590958; PIDN:AAB98208.1; PID:g1590960

C:Genetics:

A:Map position: REV214151-212064

A:Start codon: GTG

C:Keywords: hydrolase

Query Match 10.3%; Score 70; DB 2; Length 695;  
Best Local Similarity 23.6%; Pred. No. 21;  
Matches 29; Conservative 17; Mismatches 41; Indels 36; Gaps 3;

QY 13 TLPLETPAEVNIISPETRAQAERLAIRAOLKREYLLOYNDPNRGLIENPALLRWARTI- 74  
DB 217 TLKEKLE---NVLSEIRKPEPERYDIDVECTPSSEALSKESELAISERNSLEKLLK 272  
QY 65 LRMAYARTINYPNRPPTPKNSLMGALCGFPLIYIYIKTERDREKELIOEGKLDRT 124  
DB 273 ALAQKWEK-----ELAVYELLSTIEKARGADYSGFGKTDRT 308  
QY 125 FHL 127  
DB 309 YTI 311

RESULT 10  
T18376

multidrug resistance protein 2 - malaria parasite (*Plasmodium falciparum*)

C:Species: *Plasmodium falciparum*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000

C:Accession: T18376

R:Rubio, J.P.; Cowman, A.F.

Exp. Parasitol. 79, 137-147, 1994

A:Title: Plasmodium falciparum: the pfmdr2 protein is not overexpressed in chloroquin

A:Reference number: Z18924; MUID:94333528

A:Accession: T18376

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1025 <RUB>

A:Cross-references: EMBL:U04640; NID:g439853; PID:g439854; PIDN:AAA21513.1

C:Genetics:

A:Gene: mdr2

Query Match 10.3%; Score 70; DB 2; Length 1025;  
 Best Local Similarity 33.3%; Pred. No. 35;  
 Matches 14; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 85 KNSLMGALCGFGLFIYIYTERDRKELIOGKLDRTFH 126  
 Db 500 KNSLGSVLEIGLITLYSTIKTKRKIRKANEMDNVH 541

## RESULT 11

S22440  
 Protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice  
 C:Species: Oryza sativa (rice)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
 C:Accession: S22440  
 R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuki, I.; Utsugi, M.; Gen. Genet. 233, 10-16, 1992  
 A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory  
 A:Reference number: S22440; MUID:92293101  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-294 <HAS>  
 A:Cross-references: EMBL:X60374; NID:920342; PIDN:CAA42922.1; PID:920343  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein  
 F:2-256/Domain: protein kinase homology <IN>  
 F:10-18/Region: protein kinase ATP-binding motif  
 F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.3%; Score 69.5; DB 2; Length 294;  
 Best Local Similarity 36.2%; Pred. No. 8.3;  
 Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 2 SFPKPKSSLTLPETLDPAEYN-ISPETRAQAELAIRQKREY 47  
 Db 240 AEPKQADLATIVPTLPAGLILSKMLRYEPNKRITARAQALEHEH 286

## RESULT 12

S69523  
 hypothetical protein 17 - phage HPI  
 C:Species: phage HPI  
 C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
 C:Accession: S69523  
 R:Fitelson, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Scocca, N.; Acids Res. 24, 2360-2368, 1996  
 A:Title: The complete nucleotide sequence of bacteriophage HPI DNA.  
 A:Reference number: S69503; MUID:96279738  
 A:Accession: S69523  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-298 <ESP>  
 A:Cross-references: EMBL:U24159; NID:91046235; PIDN:AA09202.1; PID:91046244  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

Query Match 10.2%; Score 69; DB 2; Length 298;  
 Best Local Similarity 26.9%; Pred. No. 9.5;  
 Matches 39; Conservative 14; Mismatches 50; Indels 42; Gaps 8;

OY 11 LRTLPETLDAEY--NISPETR---AOAERLAIRAOQKRE-YLLOYNDPGRGLIENP 63  
 Db 30 LHEAETVDEHYTANLMEPEHRFNNQGVTELKAEENKGETOLFAIIPNKE-LIEYN 88  
 OY 64 ALRWAVARTINVPNFRPPKNSL-----MGALCGFGPLI 99  
 Db 83 RAGQYLTFT-SIELTPNFRNSGKAVLSGLVTDSPASVGTTELKFNNAQSGSVCG----- 142  
 OY 100 FIYIITEDRKELIOGKLDRT 124

Db 143 ----EFIKVDSAKED-VEEEKALRT 163

## RESULT 13

S12519  
 glutactin - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000  
 C:Accession: S12519  
 R:Olson, P.F.; Fessler, L.I.; Nelson, R.E.; Sterne, R.E.; Campbell, A.G.; Fessler, J.  
 EMBO J. 9, 1219-1227, 1990  
 A:Title: Glutactin, a novel Drosophila basement membrane-related glycoprotein with se  
 A:Reference number: S12519; MUID:90214632  
 A:Accession: S12519  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1023 <OLS>  
 A:Cross-references: EMBL:X53286; NID:9297084; PIDN:CAA37380.1; PID:9297085  
 C:Genetics:  
 A:Introns: 390/3

Query Match 10.2%; Score 69; DB 2; Length 1023;  
 Best Local Similarity 26.5%; Pred. No. 44;  
 Matches 36; Conservative 15; Mismatches 45; Indels 40; Gaps 8;

OY 2 SFPKPKSSLTLPETLDPAEYNISPETRAQAE---RLAIRAOQKREYLYOY-----N 52  
 Db 790 SYEYGPENGNLEPET--DANRNFSEEDREQOQOQLRREQOQOEREYQLQREDOER 847

OY 53 DPNRGLIE-NPALLRWAVARTINVPNFR-----TPKNSLMGALCGFGLFIYIYIK 106  
 Db 848 EOERQGOEPP-----EYPSYEEYSRALQKNAERDRI-----YAE 886

OY 107 TERDR--KEKLIQCK 120  
 Db 887 QERERQOQETLQENQ 902

## RESULT 14

T18561  
 vitellogenin vit-6 [similarity] - Oscheilus sp. (strain CEM1)  
 C:Species: Oscheilus sp.  
 A:Variety: PS1131  
 C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 01-Dec-2000  
 C:Accession: T18561  
 R:Winter, C.E.; Penha, C.; Blumenthal, T.  
 Mol. Biol. Evol. 13, 674-684, 1996  
 A:Title: Comparison of a vitellogenin gene between two distantly related rhaditid ne  
 A:Reference number: Z18974; MUID:96212969  
 A:Accession: T18561  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1660 <WIN>  
 A:Cross-references: EMBL:U35449; NID:91515336; PID:91515337; PIDN:AA849749.1  
 C:Genetics:  
 A:Introns: 28/1; 100/3; 171/3; 791/1  
 A:Note: CEM1-vit-6  
 C:Superfamily: vitellogenin

Query Match 10.2%; Score 69; DB 2; Length 1660;  
 Best Local Similarity 24.8%; Pred. No. 81;  
 Matches 34; Conservative 20; Mismatches 57; Indels 26; Gaps 4;

OY 6 YKPSLTLPETLDPAEYNISPETRAQAEALIRAOQKREYLYOYNDPGRGLIENPAL 65  
 Db 631 YTWSTLKTISESNPAE---KEIRRVSSGLASTIVEEOKYLESKHKTFNFMOSGAT 686  
 OY 66 LRWAVARTINVPNFRPPKNSLMGALCGF-----PLIFY-----YIKTE 108  
 Db 687 LMAA-----TIFSDSVLPRETPASLETVEGGMNKYLAQIGLYONNDSVLSKLQKVE 741

QY 109 RDREKLIQEGKLDRTF 125  
 1:1: 11:1  
 Db 742 ETGLEQLVVRGKRSSF 758

## RESULT 15

S23095

protein kinase (EC 2.7.1.37) cdc2 - Arabidopsis thaliana

N:Alternate names: cdc2 protein homolog; CDC2a protein; cell division control protein 2

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 22-Nov-1993 #sequence\_revision 03-Nov-1995 #text\_change 20-Jun-2000

C:Accession: S23095; A48984; J01337; J00967; S18202

R:Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.  
 FEBS Lett. 304, 73-77, 1992

A:Title: Exon-intron organization of the Arabidopsis thaliana protein kinase genes CDC2a

A:Reference number: S23095; MUID:92316202

A:Accession: S23095

A:Molecule type: DNA

A:Residues: 1-294 &lt;IMA&gt;

A:Cross-references: EMBL:DI0850; NID:g217848; PIDN:BAA01623.1; PID:g217849

R:Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.  
 Biochem Soc Trans. 20, 80-84, 1992

A:Title: Control of cell division in plants.

A:Reference number: A48984; MUID:92339744

A:Accession: A48984

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-294 &lt;INZ&gt;

A:Experimental source: flower

A:Note: Sequence extracted from NCBI backbone (NCBIP:109461)

R:Hirayama, T.; Imajuku, Y.; Anal, T.; Matsui, M.; Oka, A.  
 Gene 105, 159-165, 1991

A:Title: Identification of two cell-cycle-controlling cdc2 gene homologs in Arabidopsis

A:Reference number: J01337; MUID:92039027

A:Accession: J01337

A:Molecule type: mRNA

A:Residues: 1-294 &lt;HIF&gt;

A:Cross-references: EMBL:X57839; NID:g16218; PIDN:CAA40971.1; PID:g16219

R:Ferreira, P.C.G.; Hemerly, A.S.; Villarroel, R.; Van Montagu, M.; Inze, D.  
 Plant Cell 3, 531-540, 1991

A:Title: The Arabidopsis functional homolog of the p34cdc2 protein kinase.

A:Reference number: J00967; MUID:93005715

A:Accession: J00967

A:Molecule type: mRNA

A:Residues: 1-294 &lt;FER&gt;

A:Cross-references: GB:S45387; NID:g257373; PIDN:AAB23643.1; PID:g257374

A:Experimental source: flower

C:Comment: The protein is a key component of the eukaryotic cell cycle.

C:Accession:

A:Residues: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3

A:Introns: 3/3

C:Function: phosphotransferase; protein kinase; required for G1 to S-phase transition

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein K

F:2-256/Domain: protein kinase homology &lt;KIN&gt;

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.1%; Score 68.5; DB 2; Length 294;

Best Local Similarity 36.2%; Pred.No.11;

Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 SFPKKPSSLRTPETLPDAEYN-ISPETRAQAERLAIRAOLKREY 47

Db 240 AFPKWKPTDLETFVNPDPDGVLLSKMLMDPTKRINARAALHEHY 286

Search completed: April 29, 2001, 06:45:36  
 Job time: 307 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2001, 06:45:14 ; Search time 9.22 Seconds

(without alignments)  
479.280 Million cell updates/sec

Title: US-09-726-899-3  
Perfect score: 678  
Sequence: 1 MSPPKPPSSLRTPLETLDP.....DRKELIOEGKLDRTFLSY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

T number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	673	99.3	128	1 NB5M_HUMAN
2	509	75.1	128	1 NB5M_BOVIN
3	315	46.5	133	1 NB5M_CHICK
4	88.5	13.1	1078	1 GYRB_SYNY3
5	71	10.5	270	1 FRG_ZYMO
6	70.5	10.4	771	1 DNK3_SYNY3
7	70	10.3	384	1 SX18_HUMAN
8	70	10.3	695	1 VATI_METUA
9	69.5	10.3	294	1 CC21_ORYSA
10	69	10.2	298	1 VPO_BPHPI
11	69	10.2	658	1 MTHR_HUMAN
12	69	10.2	1023	1 GUT_DROME
13	69	10.2	1660	1 VIT6_OSCBR
14	68.5	10.1	294	1 CC2A_ARATH
15	68.5	10.1	574	1 VJ23_YEAST
16	68.5	10.1	1333	1 RDPO_SCHPO
17	68	10.0	270	1 LPSC_RHIME
18	67.5	10.0	153	1 VPG_EYDVP
19	67	9.9	455	1 P2X5_RAT
20	67	9.9	618	1 M3K2_HUMAN
21	66.5	9.8	1036	1 SECA_SPIOL
22	66	9.7	193	1 LOIB_NEIMA
23	66	9.7	657	1 VP4B_FOMPV
24	66	9.7	677	1 RN14_YEAST
25	65.5	9.7	450	1 DHE4_LACBI
26	65.5	9.7	2358	1 MOKD_SCHPO
27	64.5	9.5	212	1 DSBA_BURCE
28	64.5	9.5	315	1 RSEB_HAEIN
29	64.5	9.5	654	1 P4I792_HAEMO
30	64.5	9.5	842	1 YJEF_YEAST
31	64.5	9.5	1003	1 ATC_ARTST
32	64.5	9.5	1247	1 NOS_ANOST
33	63.5	9.4	294	1 CC2_MAIZE

34	63.5	9.4	461	1 MANA_EMENT
35	63.5	9.4	686	1 KNLG_STRPU
36	63.5	9.4	878	1 SYA_BACSU
37	63	9.3	188	1 YB69_HAEIN
38	63	9.3	285	1 YN26_YEAST
39	63	9.3	400	1 YL37_CAEEL
40	63	9.3	424	1 KASA_STRCO
41	63	9.3	721	1 CLAT_DROME
42	62.5	9.2	587	1 UL84_HCMVT
43	62.5	9.2	966	1 AMPN_FELCA
44	62.5	9.2	1521	1 EMS5_CAEEL
45	62.5	9.2	1603	1 VIT5_CAEEL

## ALIGNMENTS

RESULT	ID	Sequence	Score	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
1	NB5M_HUMAN	095168: 1	673	99.3%	100.0%	128	0	0	0	0
2	NB5M_HUMAN	095168: 1	509	75.1%	100.0%	128	0	0	0	0
3	NB5M_HUMAN	095168: 1	315	46.5%	100.0%	128	0	0	0	0
4	GYRB_SYNY3	095168: 1	88.5	13.1%	100.0%	128	0	0	0	0
5	FRG_ZYMO	095168: 1	71	10.5%	100.0%	128	0	0	0	0
6	DNK3_SYNY3	095168: 1	70.5	10.4%	100.0%	128	0	0	0	0
7	SX18_HUMAN	095168: 1	70	10.3%	100.0%	128	0	0	0	0
8	VATI_METUA	095168: 1	70	10.3%	100.0%	128	0	0	0	0
9	CC21_ORYSA	095168: 1	69.5	10.3%	100.0%	128	0	0	0	0
10	VPO_BPHPI	095168: 1	69	10.2%	100.0%	128	0	0	0	0
11	MTHR_HUMAN	095168: 1	69	10.2%	100.0%	128	0	0	0	0
12	GUT_DROME	095168: 1	69	10.2%	100.0%	128	0	0	0	0
13	VIT6_OSCBR	095168: 1	69	10.2%	100.0%	128	0	0	0	0
14	CC2A_ARATH	095168: 1	68.5	10.1%	100.0%	128	0	0	0	0
15	VJ23_YEAST	095168: 1	68.5	10.1%	100.0%	128	0	0	0	0
16	RDPO_SCHPO	095168: 1	68.5	10.1%	100.0%	128	0	0	0	0
17	LPSC_RHIME	095168: 1	68	10.0%	100.0%	128	0	0	0	0
18	VPG_EYDVP	095168: 1	67.5	10.0%	100.0%	128	0	0	0	0
19	P2X5_RAT	095168: 1	67	9.9%	100.0%	128	0	0	0	0
20	M3K2_HUMAN	095168: 1	67	9.9%	100.0%	128	0	0	0	0
21	SECA_SPIOL	095168: 1	66.5	9.8%	100.0%	128	0	0	0	0
22	LOIB_NEIMA	095168: 1	66	9.7%	100.0%	128	0	0	0	0
23	VP4B_FOMPV	095168: 1	66	9.7%	100.0%	128	0	0	0	0
24	RN14_YEAST	095168: 1	66	9.7%	100.0%	128	0	0	0	0
25	DHE4_LACBI	095168: 1	65.5	9.7%	100.0%	128	0	0	0	0
26	MOKD_SCHPO	095168: 1	65.5	9.7%	100.0%	128	0	0	0	0
27	DSBA_BURCE	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
28	RSEB_HAEIN	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
29	P4I792_HAEMO	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
30	YJEF_YEAST	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
31	ATC_ARTST	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
32	NOS_ANOST	095168: 1	64.5	9.5%	100.0%	128	0	0	0	0
33	CC2_MAIZE	095168: 1	63.5	9.4%	100.0%	128	0	0	0	0

OY 122 DRTFHLSY 129  
 Db 121 DRTFHLSY 128

## RESULT 2

DB5M\_BOVIN  
 ID DB5M\_BOVIN STANDARD; PRT: 128 AA.  
 AC P48305;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE B15 SUBUNIT (EC 1.6.5.3) (EC 1.6.99.3)  
 DE (COMPLEX I-B15) (CI-B15).  
 GN NDUF84.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OA NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE-Heart;  
 RX MEDLINE=92389317; PubMed=1518044;  
 RA Walker J.E., Arizumi J.M., Dupuis A., Fearnley I.M., Finel M.,  
 Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.,  
 "Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from  
 bovine heart mitochondria. Application of a novel strategy for  
 sequencing proteins using the polymerase chain reaction."  
 J Mol. Biol. 226:1051-1072(1992).  
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
 CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
 TO BE UBIQUINONE.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: X64898; CAA46107.1;  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Acetylation.  
 FT MOD\_RES 0 0  
 FT MOD\_RES 1 1 ACETYLATION.  
 SQ SEQUENCE 128 AA; 15053 MW; CC1352E9E80DF7D5 CRC64;

Query Match  
 Best Local Similarity 75.18; Score 509; DB 1; Length 128;  
 Matches 94; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

OY 2 SPPYKPSLTLPETLDPAEVNIISPETRAOERLAIRAOLKREYLLOYNPNRGLIE 61  
 Db 1 SPPYKPSLTLPETLDPAEVNIISPETRAOERLAIRAOLKREYLLOYNPNRGLIE 61  
 OY 62 NPALLBAAVARTINVPNPPPKNSLMGALCGFPLFIYTIKTERRKREKLQEGKL 121  
 Db 61 DPALVARTINVPNPPPKNSLMGALCGFPLFIYTIKTERRKREKLQEGKL 121  
 OY 122 DRTFHLSY 129  
 Db 121 DRTFHLSY 128

RESULT 3  
 DB5M\_CHICK  
 ID DB5M\_CHICK STANDARD; PRT: 133 AA.

AC P48306;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PUTATIVE NADH-UBIQUINONE OXIDOREDUCTASE B15 SUBUNIT (EC 1.6.5.3)  
 DE (EC 1.6.99.3) (COMPLEX I-B15) (CI-B15) (HYPOTHETICAL PROTEIN WALTER)  
 DE (GGHPW).  
 GN NDUF84.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OA NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94040816; PubMed=7901127;  
 RA Goldberg G.S., Kaczmarek W.,  
 "A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1  
 homeobox is likely to encode the NADH ubiquinone oxidoreductase  
 subunit B15."  
 Gene 133:233-235(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93077061; PubMed=1359990;  
 RA Goldberg G.S., Kaczmarek W.,  
 "Sequence of a novel chicken genomic DNA fragment that hybridizes to  
 the murine Hox-3.1 homeobox."  
 Gene 121:397-398(1992).  
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
 CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
 TO BE UBIQUINONE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE  
 (BY SIMILARITY).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL: X60778; CAA43193.1; ALT\_SEQ.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 FT MOD\_RES 0 0  
 FT MOD\_RES 1 1  
 SQ SEQUENCE 133 AA; 14938 MW; ED7B82948C97B352 CRC64;

Query Match  
 Best Local Similarity 46.58; Score 315; DB 1; Length 133;  
 Matches 62; Conservative 14; Mismatches 29; Indels 0; Gaps 0;

OY 5 KYKSSLTLPETLDPAEVNIISPETRAOERLAIRAOLKREYLLOYNPNRGLIENPA 64  
 Db 12 KYKSSLTLPETLDPAEVNIISPETRAOERLAIRAOLKREYLLOYNPNRGLIENPA 64  
 OY 65 LIRWAVARTINVPNPPPKNSLMGALCGFPLFIYTIKTERRKREKLQEGKL 109  
 Db 72 LIRWAVARTINVPNPPPKNSLMGALCGFPLFIYTIKTERRKREKLQEGKL 109

RESULT 4  
 GYRB\_STNY3  
 ID GYRB\_STNY3 STANDARD; PRT: 1078 AA.  
 AC P77966;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE DNA GYRASE SUBUNIT B (EC 5.99.1.3).  
 GN GYRB OR SLT2005.  
 OS Synechocystis sp. (strain PCC 6803).



```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FORMADIDOPYRIMIDINE-DNA GLYCOSYLASE (EC 3.2.2.23) (FAPY-DNA
DE GLYCOSYLASE).
GN MUTM OR PPG.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;
OC Zymomonas.
OX NCBI_TaxID=542;
RN
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 31821 / ZM4 / CP4.
RA Lee H.J., Kang H.S.;
Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
RL
RL
CC -1- FUNCTION: THIS ENZYME MAY PLAY A SIGNIFICANT ROLE IN PROCESSES
CC LEADING TO RECOVERY FROM MUTAGENESIS AND/OR CELL DEATH BY
CC ALKYLATING AGENTS (BY SIMILARITY).
CC CATALYTIC ACTIVITY: HYDROLYSIS OF THE DEOXYRIBOSE N-GLYCOSIDIC
CC BOND TO EXCISE 2,6-DIAMINO-4'-HYDROXY-5-N-METHYLFORMADIDOPYRIMIDINE
CC (FAPY) OR 4,6-DIAMINO-5-FORMADIDOPYRIMIDINE.
CC -1- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PPG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF088986; AAD21547.1; ALT_INIT.
DR InterPro; IPR000191; -.
DR Pfam; PF01149; Fapy_DNA_glyco. 1.
DR PROSITE; PS01242; FPG; 1.
KW DNA repair; Hydrolase; Glycosidase; Zinc.
FT ZN_FING 246 268 POTENTIAL.
FT SEQUENCE 270 AA; 30693 MW; EEDDE1943D3F5F58 CRC64;
SQ
OY 42 QLRKEVLLQYNDPNRG--LIENFALLRMVYARFINYP--NFRP 82
Db 96 QTKNNEFVSLYDPRRFGSLDIYKKMQLMEWTFRNIGPEPLGTGNEP 142

RESULT 6
DNK3_SYNY3
AC P73098; STANDARD; PRT; 771 AA.
ID DNK3_SYNY3
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DNK3 PROTEIN 3 (HEAT SHOCK PROTEIN 70) (HSP70).
GN DNK3 OR SLI1932.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=97061201; Pubmed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirasawa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,
RA Tabata S.;
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
```

RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).  
 CC -1- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D90903; BAA17123.1; -  
 DR HSSP: P19120; IATS.  
 DR InterPro: IPR001023; -  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70.1; 1.  
 DR PROSITE: PS00329; HSP70.2; 1.  
 DR PROSITE: PS01036; HSP70.3; FALSE\_NEG.  
 DR Rosette: PS01036; HSP70.3; FALSE\_NEG.  
 KW Chaperone: ATP-binding; Multigene family; Heat shock.  
 SO SEQUENCE 771 AA; 86030 MW; BE715F28B372738C CRC64;  
 Query Match 10.4%; Score 70.5; DB 1; Length 771;  
 Best Local Similarity 30.5%; Pred. No. 12;  
 Matches 25; Conservative 9; Mismatches 33; Indels 15; Gaps 3;  
 Oy 16 EFTDPAEYNISPTERRAQAERLAIRALQKREYLYQNDPNRRGLIENPALLRMAYARTI- 74  
 Db 570 EILDSLEKDERLADQADQDLQVLELNEVRNQYDD-KEEGFE-----AKKTFT 621  
 Oy 75 -----NYPNPRPTPKNSLNG 90  
 Db 622 GDFDDDDYNNRRPAPRDYRG 643  
 RESULT 7  
 ID SX18\_HUMAN STANDARD; PRT; 384 AA.  
 AC P35713; Q9NPH8;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSCRIPTION FACTOR SOX-18.  
 GN SOX18.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=10858556;  
 RA Stranojic S., Stevanovic M.;  
 RT "The human SOX18 gene: cDNA cloning and high resolution mapping";  
 RL Biochim. Biophys. Acta 1492:237-241(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Azuma T., Seki N., Yoshikawa T., Masuho Y., Muramatsu M.;  
 RT "cDNA cloning, tissue expression and chromosome mapping of human  
 RL homolog of SOX18";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 96-149 FROM N.A.  
 RX MEDLINE=92310993; PubMed=1614875;  
 RA Denay P., Swift S., Brand N., Dabhadre N., Barton P., Ashworth A.;  
 RT "A conserved family of genes related to the testis determining gene,  
 RT SRY";  
 RT Nucleic Acids Res. 20:2887-2887(1992).  
 RL FUNCTION: BINDS TO THE CONSENSUS SEQUENCE 5'-AACAAAG-3' AND IS  
 CC ABLE TO TRANS-ACTIVATE TRANSCRIPTION VIA THIS SITE (BY

CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.  
 CC -1- CAUTION: WAS CALLED SOX-8 BY REF.3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ243896; CAB95835.1; -  
 DR EMBL: AB033888; BAA94874.1; -  
 DR EMBL: X65664; CAA46615.1; -  
 DR PIR: S21482; S21482.  
 DR PIR: S22941; S22941.  
 DR HSSP: Q05066; IHR2.  
 DR MIM: 601618; -  
 DR InterPro: IPR000910; -  
 DR Pfam: PF00505; HMG\_box; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein.  
 FT DNA\_BIND 85 153 HMG BOX.  
 SO SEQUENCE 384 AA; 40891 MW; 327462E519770062 CRC64;  
 Query Match 10.3%; Score 70; DB 1; Length 384;  
 Best Local Similarity 20.8%; Pred. No. 6;  
 Matches 31; Conservative 21; Mismatches 43; Indels 54; Gaps 5;  
 Oy 4 PKYPSRLTLETLDAEYNISPTERRAQAERLAIR-----AOLKREYLYQYN 52  
 Db 48 PASPPSPQSPRSPSPERGKGLSPAGRGROADESRIRPMNAFMWAKDERRLAQN 107  
 Oy 53 D-----PNRRGLIENPALLRMAYARTINYPNRRPKNSLNGAL 92  
 Db 108 PDLHNAVLKMKLGRAMKELNAEKRPFEAEERLRYOHLRD--HPNRYRPR----- 157  
 Oy 93 CGFGLPFIYIITKTERDKELIOEGKL 121  
 Db 158 -----RKKQARKARKLEPGLL 173  
 RESULT 8  
 ID VAI1\_METUA STANDARD; PRT; 695 AA.  
 AC O57675;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE V-TYPE ATP SYNTHASE SUBUNIT I (Ec 3.6.1.34) (V-TYPE ATPASE SUBUNIT I).  
 GN ATP1 OR M10222.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 NX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utechtack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith B.O., Weese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii";  
 RL Science 273:1058-1073(1996).  
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON

GRADIENT ACROSS THE MEMBRANE.  
SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U67478; AAB98208.1; -  
TIGR: MJ0222; -  
DR InterPro: IPR002490; -  
DR Pfam: PF01496; V.Arpaase\_sub\_a; 1.  
KM Hydrolase; Hydrogen ion transport; Transmembrane.  
FT TRANSMEM 392 412 POTENTIAL.  
FT TRANSMEM 425 445 POTENTIAL.  
FT TRANSMEM 497 517 POTENTIAL.  
FT TRANSMEM 534 554 POTENTIAL.  
FT TRANSMEM 556 576 POTENTIAL.  
FT TRANSMEM 605 625 POTENTIAL.  
FT TRANSMEM 627 647 POTENTIAL.  
SQ SEQUENCE 695 AA; 76953 MW; F03E5CEBEE29D53D CRC64;

Query Match 10.3%; Score 70; DB 1; Length 695;  
Best Local Similarity 23.6%; Pred. No. 12;  
Matches 29; Conservative 17; Mismatches 41; Indels 36; Gaps 3;

Y 13 TLPLETDAEYNISPETRAQAERLAI-----RAQLKREYLQYNDPNRGLIENPA 64  
D 217 TLKLEKLE-----NVLSIKKFEFERDYDISVECTPSEALSKISELSERNSLIEKLIK 272  
Y 65 LLRAVYATINYPFRFPKNSLMGALCGFPLFIYIITERTDRKKLLQEGLDLT 124  
D 273 ALAOKMER-----ELLAVALLELSIEKAKGDAYSQFGKTDRT 308  
Y 125 FHL 127  
D 309 YTI 311

RESULT 9  
CC21\_ORYSA STANDARD; PRT; 294 AA.  
AC P29618;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-).  
GN CDC2-1.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaeae; Oryza.  
OC NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV\_NIPPONBARE;  
RX MEDLINE=92293101; PubMed-1376401;  
RA Hayashino J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y., Suwaka I., Utsugi T., Toh-E A., Kikuchi Y.;  
RT "Isolation and characterization of cDNA clones encoding cdc2 homologues from Oryza sativa: a functional homologue and cognate variants".  
RL Mol. Genet. 233:10-16(1992).  
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES THE REPEITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.  
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY SIMILARITY).

SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CDC2/CDKX SUBFAMILY.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X60374; CAA42922.1; -  
DR PIR: S22440; S22440.  
DR HSHP: P24941; 1HKX.  
DR InterPro: IPR000719; -  
DR InterPro: IPR002290; -  
DR Pfam: PF00069; pkinase; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle; Cell division; Mitosis; Phosphorylation.  
FT DOMAIN 4 287 PROTEIN KINASE.  
FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
FT BINDING 33 33 ATP (BY SIMILARITY).  
FT ACT\_SITE 127 127 BY SIMILARITY.  
FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
SQ SEQUENCE 294 AA; 34071 MW; 5132293AE4FC131 CRC64;

Query Match 10.3%; Score 69.5; DB 1; Length 294;  
Best Local Similarity 36.2%; Pred. No. 5;  
Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

Y 2 SFPRKYPSSLKTLPLETDAEYN-ISPETRAQAERLAIQAOLKREY 47  
D 240 APPKQAQADLATIVPLTRAGDLDSKMLRYEPNKRITARQALEHEX 286

RESULT 10  
VPO\_BPHP1 STANDARD; PRT; 298 AA.  
AC P51719;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE PROBABLE CAPSID SCAFFOLDING PROTEIN (ORF17).  
OS Bacteriophage HPI.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.  
OC NCBI\_TaxID=10690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPIc1;  
RX MEDLINE=96279738; PubMed-8710508;  
RA Esposito D., Filzmaurice W.P., Benjamin R.C., Goodman S.D., Waldman A.S., Socca J.J.;  
RT "The complete nucleotide sequence of bacteriophage HPI DNA".  
RL Nucleic Acids Res. 24:2360-2368(1996).  
CC -1- SIMILARITY: STRONG, TO PHAGE P2 PROTEIN O.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U24159; AAB09202.1; -  
DR Capsid assembly.  
KW SEQUENCE 298 AA; 33702 MW; 6D284146BD700C1F CRC64;



SEVERE MENTAL RETARDATION, PERINATAL DEATH, PSYCHIATRIC DISTURBANCES, AND LATER-ONSET NEURODEGENERATIVE DISORDERS. - SIMILARITY: BELONGS TO THE METHYLENETETRAHYDROFOLATE REDUCTASE (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U09806; AAA74440.2; -

DR EMBL; A47328; CAA03053.1; -

DR EMBL; AF105987; AAD17965.1; -

DR EMBL; AF105977; AAD17965.1; JOINED.

DR EMBL; AF105978; AAD17965.1; JOINED.

DR EMBL; AF105979; AAD17965.1; JOINED.

DR EMBL; AF105980; AAD17965.1; JOINED.

DR EMBL; AF105981; AAD17965.1; JOINED.

DR EMBL; AF105982; AAD17965.1; JOINED.

DR EMBL; AF105983; AAD17965.1; JOINED.

DR EMBL; AF105984; AAD17965.1; JOINED.

DR EMBL; AF105985; AAD17965.1; JOINED.

DR EMBL; AF105986; AAD17965.1; JOINED.

DR EMBL; AJ237672; CAB41971.1; -

DR MIM; 236250; -

KM Oxidoreductase; Flavoprotein; FAD; Disease mutation;

KM Polymorphism.

FT VARIANT 51 51 R -> P (IN HOMOCYSTEINEMIA).

FT VARIANT 52 52 R -> Q (IN HOMOCYSTEINEMIA).

FT VARIANT 157 157 R -> Q (IN HOMOCYSTEINEMIA).

FT VARIANT 222 222 R -> Q (IN HOMOCYSTEINEMIA).

FT VARIANT 227 227 T -> M (IN HOMOCYSTEINEMIA).

FT VARIANT 251 251 P -> L (IN HOMOCYSTEINEMIA).

FT VARIANT 323 323 L -> P (IN HOMOCYSTEINEMIA).

FT VARIANT 324 324 N -> S (IN HOMOCYSTEINEMIA).

FT VARIANT 325 325 R -> C (IN HOMOCYSTEINEMIA).

FT VARIANT 335 335 R -> C (IN HOMOCYSTEINEMIA).

FT VARIANT 339 339 W -> G (IN HOMOCYSTEINEMIA).

FT VARIANT 357 357 R -> C (IN HOMOCYSTEINEMIA).

FT VARIANT 377 377 R -> C (IN HOMOCYSTEINEMIA).

FT VARIANT 387 387 G -> D (IN HOMOCYSTEINEMIA).

FT VARIANT 428 428 E -> A (COMMON POLYMORPHISM; THERMOSTABLE; DECREASED ACTIVITY).

FT VARIANT 572 572 P -> L (IN HOMOCYSTEINEMIA).

FT VARIANT 586 586 E -> K (IN HOMOCYSTEINEMIA).

SO SEQUENCE 656 AA; 74538 MW; A223439AC8554640 CRC64;

Query Match 10.2%; Score 69; DB 1; Length 656;  
Best Local Similarity 28.7%; Pred. No. 14;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

CC 35 ERLAIRAQLKREYLLQYNDPNRGLIENPALLRMVARTINYPNRPYKNSLMGALCG 94  
CC 458 ERLAETSLKREYLLRV---NRGIL-----TINSQPNINCKPSSD---PIVG 499  
CC

CC 95 EGP---LIF---IYIILKTRDKREKLIQ---EGKLDRTFHL 127  
CC 500 WQPSGQYVQKAYLEFFTSRETAELQLQVLRKREYLLRVNHL 540  
CC

CC RESULT 12  
CC GLT\_DROME STANDARD; PRT; 1023 AA.  
CC ID GLT\_DROME  
CC AC P33438;  
CC DT 01-FEB-1994 (Rel. 28, Created)  
CC DT 01-FEB-1994 (Rel. 28, Last sequence update)  
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)  
CC DE GLUTACTIN PRECURSOR.  
CC GN GLT.  
CC OS Drosophila melanogaster (Fruit fly).  
CC OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
CC OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
CC OC Ephydroidea; Drosophilidae; Drosophila.  
CC OX NCBI\_Taxid=7227;  
CC RN [1]  
CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
CC RP STRAIN=OREGON-R;  
CC RX MEDLINE=90214632; PubMed=2108864;  
CC RA Olson P.F., Fessler L.I., Nelson R.E., Sterne R.E., Campbell A.G.,  
CC RA Fessler J.H.;  
CC RT "Glutactin, a novel Drosophila basement membrane-related glycoprotein  
CC RT with sequence similarity to serine esterases."  
CC RL EMBO J. 9:1219-1227(1990).  
CC CC - FUNCTION: NOT KNOWN. BINDS CALCIUM IONS.  
CC CC - SUBCELLULAR LOCATION: BASEMENT MEMBRANES.  
CC CC - PPM: EXTENSIVELY O-GLYCOSYLATED AND ALSO N-GLYCOSYLATED.  
CC CC - PPM: FOUR TYROSINES ARE SULFATED.  
CC CC - SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE TYPE-B  
CC CC CARBOXYLESTERASE/LIPASE FAMILY.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; X53286; CAA37380.1; -

DR HSSP; P21836; IMAH.

DR FlyBase; FBgn0001114; GLT.

DR InterPro; IPR002018; -

DR Pfam; PF00135; Coesterase; 2.

DR PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; 1.

KM Glycoprotein; Sulfatation; Calcium-binding; Signal.

FT SIGNAL 1 17  
FT CHAIN 18 1023  
FT SIMILAR 18 602  
FT DOMAIN 603 615  
FT DOMAIN 616 1023  
FT CARBOHYD 115 115  
FT CARBOHYD 368 368  
FT CARBOHYD 402 402  
FT CARBOHYD 810 810  
FT DISULFID 123 145  
FT DISULFID 298 316  
SO SEQUENCE 1023 AA; 118412 MW; 3638CF79AB860E8C CRC64;

Query Match 10.2%; Score 69; DB 1; Length 1023;  
Best Local Similarity 26.5%; Pred. No. 24;  
Matches 36; Conservative 15; Mismatches 45; Indels 40; Gaps 8;  
QY 2 SPKKYKPSLSRLPTETIDPAEYINISPTERRAOE---RLAIRAQLKREYLLQY-----N 52

```

Db 790 SEQYGPGENULPET--DANKFSEEDREQOQLREGEQOERQLOLEREQER 847
OY 53 DPNRGULE-PPALLRMVARTINYPNRP-----TPKNSLMALCGGFLIFTIYIIR 106
Db 848 EOEERGOQDEPGP-----EEYPSYEYSRALOEKMERDRI-----YAE 886
OY 107 TERDR-KEKLOEGK 120
Db 887 QERERQOETILOENO 902

RESULT 13
VIT6_OSCBR STANDARD: PRT: 1660 AA.
AC 094637:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VITELLOGENIN 6 PRECURSOR.
GN VIT-6.
OS Oscheilus brevesophaga.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Rhabditinae; Oscheilus.
OX NCBI_TaxID=57871;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CEM1:
RX MEDLINE=96212989; PubMed=8676742;
RA Winter C.E., Penha C., Blumenthal T.;
RT "Comparison of a vitellogenin gene between two distantly related
RL Rhabditid nematode species.";
RL Mol. Biol. Evol. 13:674-684(1996).
CC -1- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE SOURCES OF
CC NUTRIENTS DURING EMBRYONIC DEVELOPMENT (POTENTIAL).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U35449; AAB49749.1; -
DR HSSP: P11064; IPHR.
DR InterPro: IPR001747; -
DR InterPro: IPR001846; -
DR fam: PRO1347; Vitellogenin.N; 1.
DR fam: PRO0094; vwd; 1.
KW Storage protein; Multigene family; Signal.
FT CHAIN 1 15
FT SIGNAL 1 15
FT CARBOHYD 237 1660
FT CARBOHYD 383 383
FT CARBOHYD 693 693
FT CARBOHYD 1307 1307
FT CARBOHYD 1596 1596
FT CARBOHYD 1629 1629
FT CARBOHYD 192108 MW; 25782630678ABBD CMC64;
SQ SEQUENCE 1660 AA; 192108 MW; 25782630678ABBD CMC64;

Query Match 10.2%; Score 69; DB 1; Length 1660;
Best Local Similarity 24.8%; Pred. No. 43;
Matches 34; Conservative 20; Mismatches 57; Indels 26; Gaps 4;
OY 6 YKRSRLTLPETIDPAEYNISPTERRAQERLAKRQKREYLLQVNDPNRGLIENPAL 65
Db 631 YTWSTLTKTJSESENPAE---KEIRRVSSQSLASIVEEQKYLESKHKTFNMFGSCAT 686
OY 66 LHMAYARTINVPNFPNLSLMALCGF-----PLIFTI-----YIKTE 108
Db 687 LNM-----TIFSNDSVLKREITASLETFVGGMNKNYLAOIGLYQNNLDSVLKLOKVE 741

```

```

OY 109 RDRKEKLOEGKIDRPF 125
Db 742 ETGLEQLVVRGKRSSSF 758

RESULT 14
CC2A_ARATH STANDARD: PRT: 294 AA.
AC P24100;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A (EC 2.7.1.-).
GN CDC2A OR CDC2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005715; PubMed=1840925;
RA Ferreira P.C.G., Hemery A.S., Villarroel R., van Montagu M., Inze D.;
RT "The Arabidopsis functional homolog of the p34cdc2 protein kinase.";
RL Plant Cell 3:531-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=92039027; PubMed=1937013;
RA Hirayama T., Imajuku Y., Anal T., Matsui M., Oka A.;
RT "Identification of two cell-cycle-controlling cdc2 gene homologs in
RL Arabidopsis thaliana.";
RL Gene 105:159-165(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316202; PubMed=1618302;
RA Imajuku Y., Hirayama T., Endoh H., Oka A.;
RT "Exon-intron organization of the Arabidopsis thaliana protein kinase
RL genes CDC2a and CDC2b.";
RL FEBS Lett. 304:73-77(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339744; PubMed=1634002;
RA Inze D., Ferreira P.C.G., Hemery A.S., van Montagu M.;
RT "Control of cell division in plants.";
RL Biochem. Soc. Trans. 20:80-84(1992).
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDCX SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: S45387; AAB33643.1; -
DR EMBL: M59198; AAA32831.1; -
DR EMBL: D10850; BAA01623.1; -
DR EMBL: X57839; CAA40971.1; -
DR PIR: A48984; A48984.
DR PIR: J00967; J00967.
DR PIR: J01337; J01337.
DR PIR: S23095; S23095.
DR HSSP: P24941; IHCK.

```

DR InterPro: IPR000719; -  
 DR InterPro: IPR002290; -  
 DR Pfam: PF00069; Kinase; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 4 287  
 FT NP\_BIND 10 18  
 FT BINDING 33 33  
 FT ACT\_SITE 127 127  
 FT MOD\_RES 14 14  
 FT MOD\_RES 15 15  
 FT MOD\_RES 161 161  
 SO SEQUENCE 294 AA; 34030 MW; B5FAE5FA9EC366E CRC64;  
 BY SIMILARITY.  
 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).

Query Match 10.1%; Score 68.5; DB 1; Length 294;  
 Best Local Similarity 36.2%; Pred. No. 6.3;  
 Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;  
 OY 2 SEPKYKPSLTLPETLPAEYN-ISPETRAQARLAIKRLKREY 47  
 DB 240 AFPKWKPTDLEFVNPDDGVLLSKMLMDPTKRINARAALHEHY 286

RESULT 15  
 YJ23\_YEAST  
 ID YJ23\_YEAST STANDARD: PRT; 574 AA.  
 AC P47113;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHEICAL. 66.1 KDA PROTEIN IN RAD7-HIT1 INTERGENIC REGION.  
 GN YJ0053W OR J1667.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C;  
 RX MEDLINE=95397595; PubMed=7668047;  
 RA Huang M.-E., Chuat J.-C., Galibert F.;  
 RT "Analysis of a 42.5 kb DNA sequence of chromosome X reveals three  
 RT tRNA genes and 14 new open reading frames including a gene most  
 RT probably belonging to the family of ubiquitin-protein ligases.";  
 RT Yeast 11:775-781(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L36344; AAA88756.1; -  
 DR EMBL: Z49553; CA89581.1; -  
 KW Hypothetical protein.  
 FT DOMAIN 193 201  
 FT DOMAIN 265 268  
 FT POLY-SER.  
 SO SEQUENCE 574 AA; 66086 MW; 1EA75D26FCF0A1DA CRC64;

Query Match 10.1%; Score 68.5; DB 1; Length 574;  
 Best Local Similarity 22.4%; Pred. No. 14;  
 Matches 22; Conservative 20; Mismatches 23; Indels 33; Gaps 5;

OY 4 PKYKSSLTLPETLPAEYN-----SPTTRAQ-----AERLAIR----- 40  
 DB 259 PFLKSSSSSLPKISPAQYDIVKHDELLTGLHRRQDMNTQOELDSFKERKSVRHCSN 318

OY 41 -----AQLKR-EYLLQYNDPNRRG-LIENPALLRW 68  
 DB 319 QNVQLNGPAKIKITIKQIDHNTMPMKKGSMLYNPKTKWK 356

Search completed: April 29, 2001, 06:47:26  
 Job time: 132 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2001, 06:45:39 ; Search time 19.67 Seconds

(without alignments)  
768.674 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678  
Sequence: 1 MSFPRKPSRLTLPETLPD.....DRKEKLIQEGKIDRTFHLISY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

T number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

- 1: SPREMBL\_15:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_mammal:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_todent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.5	12.1	113	5	Q9V753 drosophila
2	82	12.1	386	2	Q9L7P6 yersinia pe
3	74.5	11.0	313	2	Q9K198 bacteroides
4	74	10.9	195	11	Q9MUN8 mus musculu
5	74	10.9	451	1	Q9V150 pyrococcus
6	74	10.9	526	2	Q9WXR9 lactobacill
7	70.5	10.4	520	10	Q9LFE21 arabidopsis
8	70.5	10.4	585	10	Q9LFG30 arabidopsis
9	70.5	10.4	2067	5	Q9U3U8 plasmodium
10	70	10.3	384	4	Q9NPH8 homo sapien
11	70	10.3	1025	5	Q25693 plasmodium
12	69.5	10.3	280	10	Q9XRF46 phaseolus a
13	69.5	10.3	280	10	Q9XFL3 phaseolus v
14	69.5	10.3	742	2	Q9ZB47 streptococc
15	69.5	10.3	1276	13	Q918D1 gallus gall
16	69	10.2	679	4	Q9UGR2 homo sapien
17	69	10.2	759	4	Q9UGR8 homo sapien
18	69	10.2	760	4	Q9Y380 homo sapien
19	69	10.2	1026	5	Q9V1J3 drosophila

20	68.5	10.1	294	10	Q9M307 arabidopsis
21	68.5	10.1	921	10	Q9SKT6 arabidopsis
22	68.5	10.1	1330	3	Q01910 schizosacch
23	68.5	10.1	1333	3	Q9UR07 schizosacch
24	68	10.0	333	4	Q9Y220 homo sapien
25	68	10.0	459	3	Q9P8W1 candida alb
26	68	10.0	523	2	Q9R0S5 mycoplasma
27	68	10.0	526	2	Q9R0S7 mycoplasma
28	68	10.0	716	11	P70521 rattus norv
29	68	10.0	756	2	Q9K1R1 serralia ma
30	67.5	10.0	289	2	Q9RUB0 delnoccocus
31	67.5	10.0	296	2	Q9PCP8 xylella fas
32	67	9.9	194	2	Q9KRP4 vibrio chol
33	67	9.9	232	1	Q9V0M0 pyrococcus
34	67	9.9	316	5	Q17824 caenorhabd
35	67	9.9	319	9	Q9MCM2 streptococ
36	66.5	9.8	153	14	Q93203 barley yell
37	66.5	9.8	453	3	Q9P4E6 candida alb
38	66.5	9.8	472	1	Q49179 methanother
39	66.5	9.8	472	2	Q9L8K2 rickettsia
40	66.5	9.8	619	4	Q9NYK3 homo sapien
41	66.5	9.8	3010	14	P90192 hepatitis c
42	66	9.7	271	5	Q22543 caenorhabd
43	66	9.7	342	2	Q83458 treponema p
44	66	9.7	406	10	Q9SLX9 psittolum nu
45	66	9.7	413	2	Q9RTK1 delnoccocus

## ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	113 AA.
Q9V753	Q9V753	01-MAY-2000 (TREMBlrel. 13, Created)		
AC	Q9V753	01-MAY-2000 (TREMBlrel. 13, last sequence update)		
DT	01-MAY-2000 (TREMBlrel. 13, last sequence update)			
DT	01-MAY-2000 (TREMBlrel. 13, last sequence update)			
DE	CG12859 PROTEIN.			
GN	CG12859			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,			
RA	Abriil J.F., Adavani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Bayendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merulov G., Mishina N.V., Mobarly C., Morris J., Moshrefi A.,			

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtak R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang X., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003813; AAF58210.1;--  
 DR FLYBASE: FBgn0033961; CG12859.  
 SQ SEQUENCE 113 AA; 13134 MW; 25A8CDB90F2DB414 CRC64;

Query Match 14.2%; Score 96.5; DB 5; Length 113;  
 Local Similarity 29.3%; Pred. No. 0.01; Mismatches 36; Indels 9; Gaps 3;  
 Matches 27; Conservative 20;

OY 42 QLRKREYLQNDPNRR-----GLIENPALLRMAYARTINVPNRPKNSLGMALCGFG 96  
 DB 20 KLRQELKSSNPRIATGEGTVPDAGLARQAMVSN-YEHFKPTGKFRIGLFAVVL 78  
 OY 97 PLFIYIYIKTERDRREKLIQESKL--DRTF 125  
 DB 79 PIALYAMALKARDGREGREKYRTGOVAVKDRDF 110

RESULT 2  
 Q9L7P6 PRELIMINARY; PRT; 386 AA.  
 AC Q9L7P6;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE PARC (FRAGMENT).  
 GN PARC.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_Taxid=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM;  
 RA Lindler L.E., Jahan N.;  
 RT Non-radioactive detection of ciprofloxacin resistance in *Yersinia*  
 pestis.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF221695; AAF36389.1; --  
 FT NON\_TER 1  
 FT TER 386  
 SQ SEQUENCE 386 AA; 43477 MW; CC748E6024664CD0 CRC64;

Query Match 12.1%; Score 82; DB 2; Length 386;  
 Best Local Similarity 27.7%; Pred. No. 1.5; Mismatches 59; Indels 18; Gaps 5;  
 Matches 36; Conservative 17;

OY 11 LRLPPTLPDAEYNISPEERRAQAER-----LAIRAQLKREYLQY-----NDPNRGLI 60  
 DB 252 LRDESQHEHNTRLVIVPRTNRVLDQVMHLFTTDERSYRINMNMIGLDNRPSVAGLL 311  
 OY 61 ENNALLRMAYAR--TINVYNFRPTP-----KNSLGMALCGFGPLFIYIYIKTERDRK 114  
 DB 312 E--ILTEWLFRROTVNRLNRLKLEKVLKRLHILEGLLFLNIDEVIIHITREDDKPL 369

OY- 115 LLOEGKLDRT 124  
 DB 370 LMDRESISFT 379

RESULT 3  
 ID Q9K198 PRELIMINARY; PRT; 313 AA.  
 AC Q9K198;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE HYPOHETICAL 37.1 KDA PROTEIN.  
 OS Bacteroides uniformis.  
 OC Bacteria; CFB group; Bacteroidaceae; Bacteroides.  
 OX NCBI\_Taxid=820;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115334; PubMed=10648516;  
 RA Shoemaker N.B., Wang G.R., Salyers A.A.;  
 RT "Multiple gene products and sequences required for excision of the  
 RT mobilizable integrated Bacteroides element NB01.";  
 RL J. Bacteriol. 182:928-936(2000).  
 DR EMBL: AP238307; AAF74439.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 313 AA; 37128 MW; DID43577FDB882D9 CRC64;

Query Match 11.0%; Score 74.5; DB 2; Length 313;  
 Best Local Similarity 22.4%; Pred. No. 6.9; Mismatches 19; Indels 11; Gaps 3;  
 Matches 15; Conservative 22;

OY 4 PKYPSRLRLPETLPDAEYNISPEERRAQAERLAIRAQLKREYLQYNDPNRGLIENP 63  
 DB 51 PKYEAARKRIGDDRRNP-DFE-----ELENVLLQSLAEKIVNEHPTNKRISIV--P 99

OY 64 ALRMAY 70  
 DB 100 AMLAMLY 106  
 RESULT 4  
 ID Q9WUN8 PRELIMINARY; PRT; 195 AA.  
 AC Q9WUN8;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE LADYBIRD-LIKE HOMEDOMAIN PROTEIN LBX2.  
 GN LBX2H OR LBX2.  
 GN Mus musculus (Mouse).  
 OS Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chen F., Liu K.C., Epstein J.A.;  
 RT "Lbx2, a novel murine homobox gene related to the *Drosophila* ladybird  
 RT genes is expressed in the developing urogenital system, eye, and  
 RT brain.";  
 RL Mech. Dev. 0:0-0(1999).

RL EMBL: AF146150; AAD31905.1; --  
 DR HSSP: P23441; 1FTT.  
 DR MGD: MGI:1342288; Lbx2h.  
 DR INTERPRO: IPR000047; --  
 DR INTERPRO: IPR001356; --  
 DR PIRAM: PF000046; homobox; 1.  
 DR PRINTS: PR00024; HOMEOBOX.  
 DR PRINTS: PR00031; HTHREPRESSR.  
 DR PROSITE: PSS0071; HOMEOBOX\_2; 1.  
 KW Homobox; DNA-binding; Nuclear protein.  
 SQ SEQUENCE 195 AA; 20916 MW; E106CE108933DB45 CRC64;

Query Match 10.9%; Score 74; DB 11; Length 195;  
 Best Local Similarity 29.2%; Pred. No. 4.4;

Matches 33; Conservative 15; Mismatches 39; Indels 26; Gaps 7;

Query Match 10.9%; Score 74; DB 1; Length 451;  
 Best Local Similarity 32.6%; Pred. No. 12;  
 Matches 29; Conservative 10; Mismatches 26; Indels 24; Gaps 5;

QY 11 LRLTPELTPDAEYNISPE-----TRAOERLAIRAO-----LKREYLQ-YNDNR 56  
 DB 57 SPRTPTQ---PSEORAPAPAPGACVRRRRKRTATTAQVLELEKRFYFOKTLASE 113

OY 57 R-GLIENPALIRMAVARTINYPNFRPTPKNSL-----MGALCGFGLIFITY 102  
 DB 114 RDGL---AARGLANAOVTFQNRRAKLRDVEEMRADVASLGLSPGVLCY 163

RESULT 5  
 OYV150 PRELIMINARY; PRT; 451 AA.

AC 09V150;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE TRYPTOPHAN SYNTHASE, SUBUNIT BETA (TRPB-2).  
 GN -PAB1970.  
 OS Pyrococcus abyssi.  
 OC Archaea: Euryarchaeota: Thermococcales: Thermococcaceae: Pyrococcus.  
 NCBI\_TaxID=29292;  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ORSAY;  
 RA Heilig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ248284; CAB49501.1;  
 DR HSSP: P00933; 2WSY.  
 DR INTERPRO: IPR001926;  
 DR PRAM: PF00291; PALP.1;  
 DR SEQUENCE 451 AA; 50297 MW; C496F6E2DFE67289 CRC64;

Query Match 10.9%; Score 74; DB 1; Length 451;  
 Best Local Similarity 32.6%; Pred. No. 12;  
 Matches 29; Conservative 10; Mismatches 26; Indels 24; Gaps 5;

QY 11 LRLTPELTPDAEYNISPE-----TRAOERLAIRAO-----LKREYLQ-YNDNR 56  
 DB 29 LPDPELTPDP---LDPEETPIDIEKLRIFAEL-VKQETSR---RY-----IE 73

OY 62 NPALLRMAVARTINYPNFRPTPKNSL 90  
 DB 74 IPGELRLKLYSKIGRPTPLFRATNLEKLG 102

RESULT 6  
 OYV150 PRELIMINARY; PRT; 526 AA.

AC 09V150;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 DE BETA N-ACETYLGLUCOSAMINIDASE.  
 OS Lactobacillus casei.  
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Lactobacillaceae; Lactobacillus.  
 NCBI\_TaxID=1582;  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 27092;  
 RA Senba M., Nakashima Y., Miake F., Watanabe K.;  
 RT Cloning and Expression of the beta-N-Acetylglucosaminidase Gene from Lactobacillus casei ATCC 27092 and Characteristics of the enzyme expressed in Escherichia coli.  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB025100; BAA76352.1;  
 DR SEQUENCE 526 AA; 56835 MW; 255D541FA300C723 CRC64;

Query Match 10.9%; Score 74; DB 2; Length 526;  
 Best Local Similarity 30.0%; Pred. No. 15;  
 Matches 27; Conservative 17; Mismatches 36; Indels 10; Gaps 4;

OY 10 SLRTLPETLPAEYNISPE-----TRAOERLAIRAO-----LKREYLQ-YNDNR 67  
 DB 60 ALRLT---LDDMSVTINGTGSVDVKAQIQGNK---AYYDDPGNALITQDMQAVLK 112

OY 68 WARTINYPNFR-PTPKNSL 96  
 DB 113 YAAARDINIIPIVNSPGHMDALITMAQLG 142

RESULT 7  
 OYV150 PRELIMINARY; PRT; 520 AA.

AC 09V150;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE F20N2.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta; Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II;  
 OC Brassicales: Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C., Shin P., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Tortum M., Vaysberg M., Yu G., Federspiel N.A., Theologis A., Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome 1."  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y., Walker M.M., Altafi H., Araujo R.R., Conn L.L., Conway A.A.B., Gonzalez A.A., Hansen N.N.F., Huijar L.L., Kremenetska I.I., Lenz C.C., Li J.J., Liu S.S., Luos S.S., Rowley D.D., Schwartz J.J., Tortum M.M., Vysotska V.V., Yu G.G., Davis R.R.W., Federspiel N.N.A., Theologis A.A., Ecker J.J.R.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN NCBI  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharly N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Tortum M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC002328; AAF79499.1;  
 DR SEQUENCE 520 AA; 60115 MW; 3945B20167DCC56B CRC64;

RE SEQUENCE FROM N.A.

SEQUENCE FROM N.A.  
[1]  
RN RP

Query Match	10.38;	Score 70;	DB 5;	Length 1025.
Best Local Similarity	33.38;	Pred. No. 86;		

ID	Q9XE13	PRELIMINARY;	PRT;	280 AA.
AC	Q9XE13			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)			
DE	CELL DIVISION CONTROL PROTEIN 2 (FRAGMENT).			
OS	Phaseolus vulgaris (Kidney bean) (French bean).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;			
OC	Fabales; Fabaceae; Papilionoideae; Phaseolus.			
OX	NCBI_TaxID=3885;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RP	Luo S., Liu S.-L., Wang Y., Wang Y.-C., Han B.-W.;			
R7	"CcC2 gene cloning and its expression along with differentiation of			
RT	adventitious root in mungbean."			
RL	Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AF126737; AAC04994.1; -			
DR	INTERPRO: IPR000719; -			
DR	INTERPRO: IPR001245; -			
DR	INTERPRO: IPR002290; -			
DR	PFAM: PF00069; pkinase. 1.			









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2001, 06:34:44 ; Search time 20.28 Seconds

(Without alignments)  
363.612 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678  
Sequence: 1 MSFPRKPSRLTLPETLDP.....DRKEKIQGKLDRTFLSY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

T number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_0401:\*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:\*
- 22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	19	W69225
2	503	74.2	113	20	Y76629
3	76	11.2	16	21	B53504
4	72	10.6	290	21	G33365
5	69.5	10.3	292	21	B35797
6	69.5	10.3	294	20	W95690
7	69	10.2	656	17	R88358
8	69	10.2	660	17	R88359
9	69	10.2	660	21	Y96186
10	68.5	10.1	204	21	G34044
11	68.5	10.1	237	21	G11215

12	68.5	10.1	237	21	G54043
13	68.5	10.1	294	21	G11214
14	68.5	10.1	294	21	G54042
15	68.5	10.1	350	21	G11213
16	68	10.0	196	20	Y02827
17	68	10.0	252	21	Y73366
18	68	10.0	371	20	Y29193
19	68	10.0	864	20	Y03636
20	67	9.9	455	19	W76439
21	66	9.7	193	21	Y75683
22	66	9.7	343	21	G47999
23	64.5	9.5	185	21	Y28426
24	64	9.4	152	12	R15223
25	64	9.4	287	21	B25784
26	64	9.4	293	21	Y76168
27	64	9.4	305	20	Y29975
28	64	9.4	321	21	B42915
29	63.5	9.4	201	21	G59502
30	63.5	9.4	204	21	G33358
31	63.5	9.4	237	21	G33357
32	63.5	9.4	294	21	G33356
33	63.5	9.4	365	16	R67590
34	63.5	9.4	416	20	Y21591
35	63.5	9.4	531	21	Y75537
36	63	9.3	424	21	Y77281
37	63	9.3	424	21	Y78833
38	62.5	9.2	171	21	G54995
39	62.5	9.2	174	21	G14734
40	62.5	9.2	200	21	G56114
41	62.5	9.2	230	21	G14733
42	62.5	9.2	235	21	G39986
43	62.5	9.2	291	21	G39985
44	62.5	9.2	306	21	G14732
45	62.5	9.2	365	16	R67592

#### ALIGNMENTS

RESULT 1	
W69225	W69225 standard; Protein: 129 AA.
XX	AC W69225;
XX	DT 18-FEB-1999 (first entry)
XX	XX
DE	NADH dehydrogenase subunit NDS-2.
XX	XX
KW	NADH dehydrogenase subunit; sympathetic nervous system disease; myopathy;
KW	nicotinamide-adenine dinucleotide dehydrogenase; cancer; leukaemia;
KW	immune system disorder; neurodegenerative disease; therapy; NDS-2.
OS	Homo sapiens.
XX	XX
PN	W09831815-A2.
XX	XX
PD	23-JUL-1998.
XX	XX
PF	17-DEC-1997; 97WO-US23970.
XX	XX
PR	17-JAN-1997; 97US-0785065.
XX	XX
PA	(INCYT-) INCYTE PHARM INC.
XX	XX
PI	Bandman O, Goli SK, Hillman JL;
XX	XX
DR	WPI: 1998-414112/35.
XX	XX
DR	N-PSDB; V44787.
XX	XX
PT	Human nicotinamide-adenine dinucleotide dehydrogenase sub:units -
PT	useful for, e.g. diagnosis, treatment and prevention of cancer,
PT	myopathy, immune system disease and neurodegeneration

Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Fragment of human  
HTRM clone 2918375  
Amino acid sequenc  
Hypoxia-regulated  
Human p53 regulate  
Neisseria meningit  
Arabidopsis thalia  
Soybean branched,c  
OV-16 antigen. On  
Human secreted pro  
Human secreted pro  
Human secreted pro  
Human electron tra  
Human ORF2679  
Arabidopsis thalia  
Zea mays protein f  
Zea mays protein f  
Hepatitis C virus  
Human secreted pro  
Neisseria meningit  
Streptomyces coeli  
Amino acid sequenc  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Hepatitis C virus

XX Claim 19; Fig 2; 80pp; English.

CC This sequence represents the NADH (reduced nicotinamide-adenine  
CC dinucleotide) dehydrogenase subunit, NDS-2, of the invention. Cells  
CC containing the DNA are used to produce the recombinant subunits.  
CC Antagonists of NDS-1 (typically antisenase sequences or ribozymes) are  
CC also used to treat or prevent cancer (leukaemia and solid cancers) and immune  
CC system disorders (e.g. asthma, diabetes, rheumatoid arthritis,  
CC osteoporosis and many others). NDS-2 and NDS-4 are used to treat myopathy  
CC (e.g. ophthalmoplegia, myoclonic epilepsy and lactic acidosis), while  
CC their antagonists are used to treat cancer and disease of the sympathetic  
CC nervous system (e.g. hypertension, arrhythmia and migraine). NDS-3 is  
CC also used to treat myopathy and its antagonists to treat cancer and  
CC neurodegenerative disease (e.g. Alzheimer's, Huntington's and Parkinson's  
CC diseases, epilepsy and Down's syndrome). In all cases NDS or their  
CC antagonists may be expressed from gene therapy vectors. Ab may be used  
CC therapeutically as antagonist; as immunoassay reagent for diagnosis or  
CC monitoring such diseases; in competitive screening assays for agents that  
CC bind specifically to the subunits, and for affinity purification of the  
CC subunits from natural sources. The DNAs are useful as primers and probes  
CC for diagnosis and monitoring (including detecting predisposition to  
CC cancer); for gene mapping or identifying related sequences, while the  
CC subunits are also used to raise antibodies and to screen for specific  
CC binding agents.

SO Sequence 129 AA;

Query Match 100.0%; Score 678; DB 19; Length 129;  
Best Local Similarity 100.0%; Pred. NO. 7.4e-74;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSFPYKSSRLTLEPTDPAEYNISPEPRRAOERLAIROLKREYLLQVNDPNRGLI 60  
DB 1 msfpykssrltpeclpaeynispctrtaealraqlkreyllqyndpnrrgl 60  
OY 61 ENPALARMAYARTIWNVNFRTPKNSLMGALCGPLFIYIITKTERDRKELIOEGK 120  
DB 61 enpallrmayartlnvnpfrtpknsimgalcgplfiyilitkterdrkeliqegk 120  
OY 121 LDRTHLSY 129  
DB 121 ldrthlsy 129

RESULT 2

ID 176629 standard; Protein; 113 AA.

AC Y76629;

DE 10-APR-2000 (first entry)

XX Human ovarian tumor EST fragment encoded protein 125.

KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;

KW gene therapy; treatment.

OS Homo sapiens.

PN DE19817557-A1.

PD 21-OCT-1999.

PF 09-APR-1998; 98DE-1017557.

PR 09-APR-1998; 98DE-1017557.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

DR WPI; 1999-591920/51.

DR N-PSDB; 277502.

XX New nucleic acid sequences expressed in ovarian, and some other, cancer  
XX tissues, and derived polypeptides, for treatment of ovarian cancer and  
XX PT identification of therapeutic agents

PS Claim 25; Page 295; 310pp; German.

CC This invention describes novel nucleic acid (cDNA) sequences (A) which  
CC have anticancer activity and are highly expressed in ovarian tumor  
CC tissue (and some also in testis and breast cancer tissue). The products  
CC of the invention can be used for gene therapy. (A) are used (i) for  
CC recombinant expression of polypeptides (B) and (ii) to isolate complete  
CC genes. (B) are used (i) to identify agents suitable for treatment of  
CC ovarian cancer; (ii) directly for treating this form of cancer (including  
CC expression from gene therapy vectors) and (iii) for generation of  
CC specific antibodies. (A) are identified by assembling ESTs (expressed  
CC sequence tags) from a particular tissue type before comparison of  
CC expression patterns. This allows a significantly longer fragment of the  
CC gene to be revealed, so should reduce the number of failures associated  
CC with the fact that ESTs from different libraries may represent different  
CC parts of the same unknown gene, distorting the estimated frequency of  
CC occurrence in a particular tissue. Y76505-Y76638 represent protein  
CC fragments encoded by the human ovarian tumor CDNA library derived EST  
CC fragments represented in 277450-277572.

SO Sequence 113 AA;

Query Match 74.2%; Score 503; DB 20; Length 113;  
Best Local Similarity 99.0%; Pred. NO. 7.2e-53;  
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 34 AERLAIRAOAKREYLLQVNDPNRGLIENPALARMAYARTIWNVNFRTPKNSLMGALC 93  
DB 18 serlallrqklkreyllqyndpnrrgl1enpallrmayartlnvnpfrtpknsimgalc 77  
OY 94 GFGPLFIYIITKTERDRKELIOEGKLDPRHLSY 129  
DB 78 gfgplfiyilitkterdrkeliqegkldprhlsy 113

RESULT 3

ID B53504 standard; Protein; 16 AA.

AC B53504;

DE 09-MAR-2001 (first entry)

XX Human colon cancer antigen protein sequence SEQ ID NO:1044.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnery;

KW immunomodulatory; muscular; gynaecological; gastrointestinal;

KW nephrotropic; antineoplastic; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder.

OS Homo sapiens.

PN WO200055351-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05883.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.



PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

QY 4 PKYKPSLRLPETLDPAEYNISPETRAQERLAIRAKLKREYLLQYN--DPNRR---- 57

Db 64 sllpymtespe-lrwafvrklyilvtqlamtavaafvkvavsnffvssnagialy 122

## RESULT 5

B35797 standard; 292 AA  
ID  
XX

AC B35797  
yy

DI 23-FEB-2001 (first entry  
XX

XX involved in cell cycle regulation SEQ ID 14.

cell cycle regulation; corn; transgenic plant; cyclin; maize; soybean;  
cyclin-dependent kinase; sunflower; sorghum; canola; wheat; alfalfa;  
cotton; rice; barley; millet.

Zea mays

PN WO2000065040-A2  
XX

02-NOV-2000.

XX 2000000-0503312

**XX**

XXX  
XX  
XXXXX

XX  
DB  
WPT.: 2000-687233 / 57

XX N-PSDB; C83104

...producing transgenic plants, preferably maize







PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158233.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 25-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 204;  
Best Local Similarity 36.2%; Pred. No. 3.2; Mismatches 21; Indels 1; Gaps 1;  
Matches 17; Conservative 8;

OY 2 SPPKPKSLRTPLETPDAETN-ISPETRRRAQERLAIRAOIKREY 47  
DB 150 aipkwkpidletfvpnlbdpdydliskmlmdpkrtrlnarealehey 196

RESULT 11  
G11215  
ID G11215 standard; Protein; 237 AA.  
AC G11215;  
XX  
WT 17-QCT-2000 (first entry)

XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9849.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130801.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132867.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137328.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.



PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140354.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142927.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 18-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150565.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.

PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155133.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0158293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160768.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161922.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 237;  
Best Local Similarity 36.2%; Pred. No. 3.9;  
Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 SPKKYPSLRPLPETHPAEYN-ISPETRAQAEIRAIQAOLKREY 47  
Db 183 atpkxpkdletfvpnlidpdydliskmlmpktrinaaalehey 229

RESULT 12  
G54043  
ID G54043 standard; Protein; 237 AA.

XX G54043;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 68864.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.  
XX EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 11-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0138458.  
PR 18-JUN-1999; 99US-0138459.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0138763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 237;  
Best Local Similarity 36.2%; Pred. No. 3.9;  
Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

OY 2 SEPKYPSLRILPELIDPAEYN-ISPETRRQAERLAIRAOLKREY 47  
D 183 afpkwptdletfvpnlpdpgvdlismldmpkkrinaaraalehey 229

RESULT 13

ID G11214 standard; Protein; 294 AA.

AC G11214;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9848.

KW Protein identification: signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129844.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132485.  
PR 05-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134766.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.



PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128314.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139753.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142290.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145921.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148688.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149920.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151067.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 23-SEP-1999; 99US-0155659.  
 PR 24-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157733.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 23-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10 1%; Score 68.5; DB 21; Length 294;  
Best Local Similarity 36.2%; Pred. No. 5.1;

Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 SEPKKXSSLRLLPELIDPAEYN-ISPETRAQAERLAIRAQIKREY 47  
Db 240 atpkwptdletfvpildpdydlkskmlmpktrinaataehy 286

RESULT 15

ID G11213 standard; Protein: 350 AA.

AC G11213;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9847.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW Hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140625.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 23-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 350;  
 Best Local Similarity 36.2%; Pred. NO. 6.5;  
 Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;  
 QY 2 SFPKPPSLRTPETLDPAYN-ISPETRRQAERLIRAQKREY 47  
 Db 296 afpwxpdlletfvpndpddvdliskmllmdpkrlnraalehey 342

Search completed: April 29, 2001, 06:45:09  
 Job time: 625 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 29, 2001, 06:40:49 ; Search time 12.61 Seconds  
(without alignments)  
196.527 Million cell updates/sec

Title: US-09-726-899-3  
Perfect score: 678  
Sequence: 1 MSFPRKPSRLTLPETLDP.....DRKELIOGKIDRPHLSY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

T number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/BackfileSI.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	2	US-08-785-065-3
2	514	75.8	129	2	US-08-785-065-10
3	69.5	10.3	294	2	US-08-874-347-26
4	69.5	10.3	294	3	US-09-093-522-26
5	69	10.2	656	3	US-08-738-000-4
6	69	10.2	660	3	US-08-738-000-2
7	64	9.4	152	1	US-07-644-372-2
8	64	9.4	305	2	US-08-946-528-1
9	63	9.3	3898	4	US-08-750-717-2
10	62.5	9.2	366	2	US-08-605-106-11
11	62	9.1	214	4	US-09-587-066-6
12	61.5	9.1	60	1	US-08-370-225-32
13	61.5	9.1	60	1	US-08-461-859-32
14	61.5	9.1	60	5	PCT-US93-10069-32
15	61.5	9.1	353	2	US-08-176-620A-14
16	61.5	9.1	353	2	US-08-461-985-14
17	61.5	9.1	484	2	US-08-836-620A-13
18	61.5	9.1	485	2	US-08-836-620A-2
19	61	9.0	420	2	US-08-846-762-73
20	61	9.0	619	1	US-08-472-934-4
21	61	9.0	619	1	US-08-472-934-12
22	61	9.0	619	2	US-08-323-460A-4
23	61	9.0	619	2	US-08-461-146C-4
24	61	9.0	619	2	US-08-461-146C-12
25	61	9.0	619	3	US-08-461-145C-4
26	61	9.0	619	3	US-08-461-145C-12
27	61	9.0	915	2	US-08-480-917-2

28	61	9.0	1213	2	US-08-937-102-2	Sequence 2, Appli
29	61	9.0	1261	1	US-08-764-100-26	Sequence 26, Appl
30	60.5	8.9	379	3	US-08-622-277A-8	Sequence 8, Appl
31	59.5	8.8	362	1	US-08-464-523B-32	Sequence 32, Appl
32	59.5	8.8	362	2	US-08-440-845D-6	Sequence 6, Appl
33	59.5	8.8	362	4	US-08-868-458-6	Sequence 6, Appl
34	59.5	8.8	716	2	US-08-766-982-1	Sequence 1, Appl
35	59.5	8.8	793	3	US-08-015-985-3	Sequence 3, Appl
36	59.5	8.8	1382	3	US-09-057-570-4	Sequence 10, Appl
37	59	8.7	262	2	US-08-719-697-10	Sequence 10, Appl
38	59	8.7	325	1	US-08-142-473A-9	Sequence 9, Appl
39	59	8.7	325	1	US-08-469-203A-9	Sequence 9, Appl
40	59	8.7	325	1	US-08-469-203A-9	Sequence 9, Appl
41	59	8.7	385	2	US-08-605-106-14	Sequence 14, Appl
42	59	8.7	565	2	US-08-892-770-6	Sequence 6, Appl
43	59	8.7	571	2	US-08-892-770-5	Sequence 5, Appl
44	59	8.7	673	4	US-09-078-347A-2	Sequence 2, Appl
45	58.5	8.6	3033	1	US-07-925-695-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-785-065-3  
: Sequence 3, Application US/08785065  
: Patent No. 5814451  
: GENERAL INFORMATION:  
: APPLICANT: Bandman, Olga  
: APPLICANT: Goli, Surya K.  
: APPLICANT: Hillman, Jennifer L.  
: TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE  
: NUMBER OF SEQUENCES: 12  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSEE: Incyte Pharmaceuticals, Inc.  
: STREET: 3174 Porter Drive  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94304  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/785,065  
: FILING DATE: Herewith  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER:  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Billings, Lucy J.  
: REGISTRATION NUMBER: 36,749  
: REFERENCE/DOCKET NUMBER: PF-0187 US  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415-855-0555  
: TELEFAX: 415-845-4166  
: TELEX:  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 129 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: LIBRARY: Consensus  
: CLONE: Consensus  
: US-08-785-065-3  
Query Match 100.0% ; Score 678 ; DB 2 ; Length 129 ;

Best Local Similarity 100.0%; Pred. No. 8.7e-75;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPKRYKSSLRTPETLPDAEYNIISPTERRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60

Db 1 MSPKRYKSSLRTPETLPDAEYNIISPTERRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60

QY 61 ENPALLMAYARTINVPNFRPTPKNSLMGALCGFGLIFLYIIKTERDRKREKLIQEGK 120

Db 61 ENPALLMAYARTINVPNFRPTPKNSLMGALCGFGLIFLYIIKTERDRKREKLIQEGK 120

QY 121 LDRTEHLISY 129

Db 121 LDRTEHLISY 129

RESULT 2  
US-08-785-065-10

Sequence 10, Application US/08785065

Patent No. 5814451

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Goll, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Inocyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/785,065

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0187 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 129 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 114

US-08-785-065-10

Query Match 75.8%; Score 514; DB 2; Length 129;

Best Local Similarity 73.6%; Pred. No. 6.5e-55;

Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSPKRYKSSLRTPETLPDAEYNIISPTERRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60

Db 1 MSPKRYKSSLRTPETLPDAEYNIISPTERRAQAERLAIRAQLKREYLLQYNDPNRRGLI 60

QY 61 ENPALLMAYARTINVPNFRPTPKNSLMGALCGFGLIFLYIIKTERDRKREKLIQEGK 120

Db 61 ENPALLMAYARTINVPNFRPTPKNSLMGALCGFGLIFLYIIKTERDRKREKLIQEGK 120

QY 121 LDRTEHLISY 129

Db 121 LDRTEHLISY 129

RESULT 3  
US-08-874-347-26

Sequence 26, Application US/08874347

Patent No. 5863741

GENERAL INFORMATION:

APPLICANT: Limper, Andrew H.

APPLICANT: Leof, Edward B.

APPLICANT: Thomas, Charles F.

APPLICANT: Gustafson, Michael P.

TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C., P.A.

STREET: 60 South Sixth Street, Suite 3300

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/874,347

FILING DATE: 13-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ellinger, Mark S.

REGISTRATION NUMBER: 34,812

REFERENCE/DOCKET NUMBER: 07039/055001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-335-5070

TELEFAX: 612-288-9696

TELEX:

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-874-347-26

Query Match 10.3%; Score 69.5; DB 2; Length 294;

Best Local Similarity 36.2%; Pred. No. 1.4;

Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

QY 2 SEPKRYKSSLRTPETLPDAEYNIISPTERRAQAERLAIRAQLKREY 47

Db 240 APPKQAOPLATIVPTLPDAGLDLSKMLRYPNKRITRQALHEHY 286

RESULT 4  
US-09-093-522-26

Sequence 26, Application US/09093522

Patent No. 6015700

GENERAL INFORMATION:

APPLICANT: Limper, Andrew H.

APPLICANT: Leof, Edward B.

APPLICANT: Thomas, Charles F.

APPLICANT: Gustafson, Michael P.

TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
TITLE OF INVENTION: CARINITI  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/093,522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 294 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-26

Query Match 10.3%; Score 69.5; DB 3; Length 294;  
Best Local Similarity 36.2%; Pred. No. 1.4;  
Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 2 SEPKKPSLRLPTLDPAEYN-ISPETFRRAQERLAIRPOLKREY 47  
Db 240 AFPKQADLALIVPTLDPAGDLDSKMLRYPKRTARQALEHEY 286

US-08-738-000-4  
Sequence 4, Application US/08738000  
Patent No. 6074821  
GENERAL INFORMATION:  
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
TITLE OF INVENTION: REDUCTASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: Continental Plaza - 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,000  
FILING DATE:  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/CA95/00314  
FILING DATE: 25-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9410620.0  
FILING DATE: 26-MAY-1994  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 656 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-738-000-4

Query Match 10.2%; Score 69; DB 3; Length 656;  
Best Local Similarity 28.7%; Pred. No. 4.8;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

OY 35 ERLAIRAQKREYLLQYNDPNRGLIENPALLRMAYARTINYPNFRPKNSLMGALCG 94  
Db 458 EPLAETSLKELLRV---NRGIL-----TINSOPNNGKRPSSD---PIVG 499

OY 95 FGP---LIF--IYYIKTERDRKELIQ---EGKLDRTFHL 127  
Db 500 WPGSGYVFORAYLEFITSRTAEALLOVLKRYELRVNYHL 540

RESULT 6  
US-08-738-000-2  
Sequence 2, Application US/08738000  
Patent No. 6074821  
GENERAL INFORMATION:  
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
TITLE OF INVENTION: REDUCTASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: Continental Plaza - 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,000  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/CA95/00314  
FILING DATE: 25-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9410620.0  
FILING DATE: 26-MAY-1994  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-738-000-2

Query Match 10.2%; Score 69; DB 3; Length 660;  
Best Local Similarity 28.7%; Pred. No. 4.8;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

OY 35 ERLAIRAQKREYLLQYNDPNRGLIENPALLRMAYARTINYPNFRPKNSLMGALCG 94  
Db 458 EPLAETSLKELLRV---NRGIL-----TINSOPNNGKRPSSD---PIVG 499

Db 462 EPLAETSLKEELRV---NRGIL-----TINSOPNKGKPSD---PIVG 503  
 QY 95 FGP---LIF---IYIKTERDRKEKLIQ---EGKIDRTFHL 127  
 Db 504 MCGSGYVOKAVLEFTRSETAELLQVLKKYELRVNHL 544

RESULT 7

US-07-644-372-2  
 ; Sequence 2, Application US/07644372  
 ; Patent No. 5416009

GENERAL INFORMATION:

APPLICANT: Lazzeri, Mario E.  
 APPLICANT: Nutman, Thomas B.  
 APPLICANT: Weiss, Nikolaus  
 TITLE OF INVENTION: A DNA SEGMENT ENCODING A SPECIFIC  
 TITLE OF INVENTION: IMMUNODIAGNOSTIC ANTIGEN  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
 STREET: 1615 L. Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/644,372  
 FILING DATE: 19910123  
 CLASSIFICATION: 435  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)861-3000  
 TELEFAX: (202)861-3000

INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 152 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-644-372-2

Query Match 9.4%; Score 64; DB 1; Length 152;  
 Local Similarity 22.2%; Pred. No. 2.7;  
 hes 26; Conservative 16; Mismatches 53; Indels 22; Gaps 3;

QY 8 PSSLRTPETLDAEYN-----ISPEYRAQAERLAIRAOLKREYLLQYNDPNRRG 58  
 Db 44 PDVYSTAPTKLVVSNLTVNMGNEITFOYKNOPTKVMDEFGALYTLVMTDPDAPS 103  
 QY 59 LIENPALLMWARTINYPNFRPTPKNSLMGALCGFGLIYIYITITENDRRKKL 115  
 Db 104 R-KNPVFRFMHMLINI-----SGQNVSSGTVLSDYIRSTKRHRISL 147

RESULT 8  
 US-08-946-528-1  
 ; Sequence 1, Application US/08946528  
 ; Patent No. 5958746

GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer L.  
 APPLICANT: Bandman, Olga  
 APPLICANT: Lal, Preeti  
 TITLE OF INVENTION: ELECTRON TRANSPORT PROTEINS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Dr.

Query Match 9.4%; Score 64; DB 1; Length 152;  
 Local Similarity 22.2%; Pred. No. 2.7;  
 hes 26; Conservative 16; Mismatches 53; Indels 22; Gaps 3;

QY 8 PSSLRTPETLDAEYN-----ISPEYRAQAERLAIRAOLKREYLLQYNDPNRRG 58  
 Db 44 PDVYSTAPTKLVVSNLTVNMGNEITFOYKNOPTKVMDEFGALYTLVMTDPDAPS 103  
 QY 59 LIENPALLMWARTINYPNFRPTPKNSLMGALCGFGLIYIYITITENDRRKKL 115  
 Db 104 R-KNPVFRFMHMLINI-----SGQNVSSGTVLSDYIRSTKRHRISL 147

CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/946,528  
 FILING DATE: Filed Herewith  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:

ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0406 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-855-0555  
 TELEFAX: 650-845-4166  
 TELEX:

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 305 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: PROSNOT16  
 CLONE: 1709102  
 US-08-946-528-1

Query Match 9.4%; Score 64; DB 2; Length 305;  
 Best Local Similarity 24.3%; Pred. No. 6.9;  
 Matches 34; Conservative 15; Mismatches 39; Indels 52; Gaps 7;

QY 8 PSSLRTPETLDAEYNISPEYRAQAERLA-----IRAOLK----- 44  
 Db 149 PSLGLTY---TGKGFHFIQPKKSPREPRVAKKLGMLAGTGITPMQLIRAILKVEDP 205  
 QY 45 -----REYL--IQYNDPNRRGL--IENPALLMWARTINYPNFRPTPKNSLMGALCGFGLI 99  
 Db 206 TQCFLEFANQTEKILILEDLELQARYPNFRKLMFTLDHP--KDMAYSKGFYADMIRE 264  
 QY 80 FRPTPKNSLMGALCGFGLI 99  
 Db 265 HLPAPGDDVILLGCGPPMV 284

RESULT 9  
 US-08-750-717-2  
 ; Sequence 2, Application US/08750717  
 ; Patent No. 6180109

GENERAL INFORMATION:  
 APPLICANT: MOORMANN, Robertus J. M.  
 APPLICANT: VAN RIJN, Petrus A.  
 TITLE OF INVENTION: Nucleotide Sequences of Pestivirus  
 TITLE OF INVENTION: Strains, Polypeptides Encoded by These Sequences and Use  
 TITLE OF INVENTION: Thereof for Diagnosis and Prevention of Pestivirus  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: YOUNG & THOMPSON  
 STREET: 745 South 23rd Street  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22202

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

Query Match 9.4%; Score 64; DB 2; Length 305;  
 Best Local Similarity 24.3%; Pred. No. 6.9;  
 Matches 34; Conservative 15; Mismatches 39; Indels 52; Gaps 7;

QY 8 PSSLRTPETLDAEYNISPEYRAQAERLA-----IRAOLK----- 44  
 Db 149 PSLGLTY---TGKGFHFIQPKKSPREPRVAKKLGMLAGTGITPMQLIRAILKVEDP 205  
 QY 45 -----REYL--IQYNDPNRRGL--IENPALLMWARTINYPNFRPTPKNSLMGALCGFGLI 99  
 Db 206 TQCFLEFANQTEKILILEDLELQARYPNFRKLMFTLDHP--KDMAYSKGFYADMIRE 264  
 QY 80 FRPTPKNSLMGALCGFGLI 99  
 Db 265 HLPAPGDDVILLGCGPPMV 284

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,717  
FILING DATE: 24-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94201743.5  
FILING DATE: 17-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/NL95/00214  
FILING DATE: 16-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: BO 39123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-521-2297  
TELEFAX: 703-685-0573  
TELEX: 248425 EMBON  
FORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3898 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-750-717-2

Query Match 9.3%; Score 63; DB 4; Length 3898;  
Best Local Similarity 23.0%; Pred. No. 2.8e+02;  
Matches 32; Conservative 19; Mismatches 44; Indels 44; Gaps 6;

QY 2 SFPRKSSSLRTLPETLDPDAEYNIISPETRAQAERLAIRAOQKREYL-LQINDPFRRLI 60  
DB 2214 SYNYFFLLNARKLDDVPPVY-----ATEDEDLAV-----ELGLDMPDPNCGTV 2260  
QY 61 ENALLMAVARTINVPNFRPPKSLMGALGFG-----PLIFIIYIKTTRD 110  
DB 2261 ETGRALQOVVGLS-----TAEMLLVAFGVYQALSKRHIPVDTIYSIEDHR- 2310  
QY 111 RKEKLIOEGLDRTFHLIS 129  
DB 2311 -----LEDTHLQY 2319

RESULT 10  
US-08-750-717-2  
Sequence 11, Application US/08605106  
Patent No. 5910631  
GENERAL INFORMATION:  
APPLICANT: Topfer, R.  
APPLICANT: Martini, N.  
APPLICANT: Scheill, J.  
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,106  
FILING DATE: 23-SEPT-1996  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02935  
FILING DATE: 01-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D.  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 235.001U51  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
FORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 366 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-605-106-11

Query Match 9.2%; Score 62.5; DB 2; Length 366;  
Best Local Similarity 36.1%; Pred. No. 13;  
Matches 13; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY 1 MSFPRKSSSLRTLPETLDPDAEYNIISPETRAQAER 35  
DB 225 LAFPEENNSLKIRKLEDDPAQYSMLEIKRRRDD 260

RESULT 11  
US-09-587-066-6  
Sequence 6, Application US/09587066  
Patent No. 6210945  
GENERAL INFORMATION:  
APPLICANT: LUNNEN, KEITH D.  
APPLICANT: MORGAN, RICHARD D.  
APPLICANT: MEISEL, TIMOTHY  
APPLICANT: WILSON, GEOFFREY G.  
TITLE OF INVENTION: METHOD FOR CLONING AND PRODUCING THE RsaI RESTRICTION  
TITLE OF INVENTION: ENDONUCLEASE IN E. COLI AND PURIFICATION OF THE  
FILE REFERENCE: NEB-179  
CURRENT APPLICATION NUMBER: US/09/587,066  
CURRENT FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 214  
TYPE: PRT  
ORGANISM: rhodopseudomonas sphaeroides  
US-09-587-066-6

Query Match 9.1%; Score 62; DB 4; Length 214;  
Best Local Similarity 29.1%; Pred. No. 7.5;  
Matches 23; Conservative 10; Mismatches 38; Indels 8; Gaps 3;

QY 11 LRTLPETLDPDAEYNIISPETRAQAERLAIRA-----QLKREYLLQYNDPFRRLIENPA 64  
DB 83 IATIPPHL-PANNSIIVARAGYPRDLFVSGATRHCEFKATSNMGGDPPRRVLTSPAT 141  
QY 65 -LLRMAVARTINVPNFRP 82  
DB 142 KMIRLVNSRQGVAPNHPV 160

RESULT 12  
US-08-370-225-32  
Sequence 32, Application US/08370225  
Patent No. 5580736  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger

TELEX: 200154



